National Defense
Defence nationale

UNCLASSIFIED

DRES

SUFFIELD REPORT

No. 525

UNLIMITED
DISTRIBUTION

3

THE SHAPE OF THE PROBABILITY DENSITY FUNCTION OF SHORT-TERM CONCENTRATION FLUCTUATIONS OF PLUMES IN THE ATMOSPHERIC BOUNDRY LAYER (U)

by

E. Yee

PCN No. 051SP

July 1989

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD : RALSTON : ALBERTA

Canada

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

RALSTON ALBERTA

SUFFIELD REPORT NO. 525

THE SHAPE OF THE PROBABILITY DENSITY FUNCTION OF SHORT-TERM CONCENTRATION FLUCTUATIONS OF PLUMES IN THE ATMOSPHERIC BOUNDARY LAYER (U)

by

E. Yee

PCN No. 051SP

ACKNOWLEDGEMENTS

The author is indebted to Dr. D. J. Ride for supplying the ion plume concentration fluctuation data. The author is grateful to Mr. S. Mellsen and Dr. S. J. Armour for valuable comments on a draft of this paper.

ABSTRACT

The shape of the probability distribution of a set of high-resolution concentration fluctuation measurements from an ion plume is studied using order statistics and certain selected quantiles derived from them. A number of graphical techniques based on the order statistics are shown to be useful for the assessment of the symmetry and tailweight of the underlying distribution of concentration. These graphical techniques are applied, from both a descriptive and a computational point of view, to elucidate the underlying distributional shape of concentration and to assess the characterization efficacy of the probability distributions that have been proposed as models for concentration fluctuations. In this respect, a new probability distribution, namely, the g and h distribution, is introduced for describing concentration fluctuations and it is shown that this distribution is superior to the more commonly used models, namely, the log-normal, the exponential, and the clipped-normal distributions utilized by previous investigators. Except for the g and h distribution, it is found that none of the commonly used models for the concentration probability distribution is able to accurately characterize the extreme upper end of the concentration frequency distribution (i.e., the end of the distribution that is critical for the prediction of the probability of exposure to peak levels). However, the clipped-normal distribution is shown to provide a reasonably conservative model for the prediction of the exceedances of critical concentration levels. Finally, it is noted that the g and h distribution yields a bimodal form for the total probability density function for concentration whereas the clipped-normal distribution provides a unimodal form. It is shown that the bimodal form of the total concentration probability density function is consistent with both the data and certain theoretical results.

INTRODUCTION

The dispersion of gaseous and particulate material released into the atmosphere is dependent on the natural mixing processes associated with the entire spectrum of turbulent and eddying motions in the atmospheric boundary layer, ranging from the microscale (the smallest being the random motion of molecules which cause molecular diffusion) up to the macroscale (e.g., arising from synoptic events and zonal currents). The eddying motion of different scales produce fluctuations of the instantaneous concentration field around the ensemble mean concentration of any contaminant injected into the turbulent flow. Consequently, the instantaneous concentration field of any contaminant released into the atmosphere is inherently stochastic in nature with the probabilistic and statistical characteristics of the field determined by the turbulent motions of the boundary flow. Typically, the root-mean-square value of concentration fluctuations is at least as large as the mean concentration and, hence, cannot be neglected in models of the turbulent dispersion of contaminant material. The problem of the statistical characterization of concentration fluctuations is of considerable interest and importance in a number of contexts including the evaluation of the screening efficacy of battlefield obscurants released into the turbulent atmosphere, the prediction of the effects of temperature fluctuations (which engender variations in the refractive index) on electromagnetic scattering and their impact on the design of electro-optical sensors for detection and passive surveillance and imaging, and the assessment of risk/hazard from the release of flammable, chemically reactive (e.g., fuel-air mixtures, liquified natural gas spills, etc.) or toxic (e.g., various chemical and biological warfare agents) materials into the atmosphere.

The statistical description of the characteristic fluctuations in concentration is conveniently embodied in the probability density function (PDF) of concentration. Indeed, in order to study and model the fluctuating concentration phenomena, it is necessary to know the probability density function or, equivalently, all the higher-order moments of the random process. With regard to the probability density function of the instantaneous concentration of a contaminant released into the atmosphere, relatively few studies have been performed—rather, the vast majority of research on atmospheric dispersion has focussed on the description of the characteristics of the ensemble mean concentration [1]. Chatwin [2] discusses the importance of the probability density function of concentration with reference to the assessment of hazards posed by the release of flammable and toxic gases; in particular, he emphasizes the need to characterize precisely the upper end or tail of the probability density function in order to determine exceedances of critical levels by peak concentrations. The shape and form of the probability density function of concentration fluctuations of dispersing clouds and plumes of contaminant in the atmospheric boundary layer have been studied by several investigators and a number of models have been proposed for the probability density function. Three commonly used probability distributions for representing concentration fluctuations are the following: (i) the log-normal distribution favored by Csanady [3] and used by Jones [4], (ii) the exponential distribution advocated by Barry [5] and Hanna [6], and (iii) the clipped-normal distribution proposed by Lewellen and

Sykes [7] through application of the principle of maximum entropy and adopted by Sawford [8], Ride [9] and Dinar *et. al* [10] as the distribution that best represents the concentration fluctuation data.

It is important to note that all the preceding probability distributions for concentration in turbulent dispersion only make use of the information embodied in the first- and second-order moments, namely, the mean concentration and the root-mean-square fluctuation in concentration, and, as such, may not adequately specify and characterize all the information embodied in the concentration fluctuations. The determination of the correct shape of the PDF of concentration may require the investigation and incorporation of the higher-order moments of the instantaneous concentration. Indeed, the higher-order moments embody the information concerning the contributions from the tails of the PDF and, in particular, these statistical parameters provide indicators of the departure of the PDF of concentration from a Gaussian form. In this context, it should be mentioned that Jones [4] and Sawford [8] have analyzed the skewness and the kurtosis measures which represent, respectively, the degree of asymmetry in the distribution and the relative peakedness or flatness in the distribution. However, these investigators have used these measures primarily as descriptive statistics and, as such, no attempts have been made to incorporate this information explicitly in models for the probability distribution of concentration. However, the probability distributions of concentration in puffs and plumes diffusing in the atmospheric boundary layer are generally non-Gaussian, undoubtedly the result of the complex nonlinear dynamical mechanisms responsible for turbulence which foster nonlinear transformations of Gaussian processes. Consequently, from this perspective, the physical nature of the stochastic processes responsible for turbulent concentration fluctuations can be only understood if higher-order moments (even beyond the 4th order) are studied with the objective of analyzing the departure of concentration statistics from Gaussian statistics.

The primary purpose of this paper is to characterize numerically the shape of the one-point probability distribution of concentration for instantaneous plumes in the atmosphere. As previously mentioned, the information about the shape and form of the PDF (especially at the lower and upper ends of the distribution) is contained in the higher-order moments of the concentration fluctuations (e.g., super- and hyper-skewness and super- and hyper-kurtosis). However, it is important to remark that since these statistical parameters depend on higher powers of the concentration data, they are generally difficult to estimate since the corresponding sample moments are very sensitive to the presence of outliers in the data. These outliers, which are simply anamolous points in the data sequence, can be due to brief instrumental malfunctions (e.g., a power surge or flicker during a point's measurement) and/or to the presence of certain uncontrollable external influences (e.g., the presence of a stray electrically charged aerosol particle). Furthermore, the outliers are frequently difficult to identify and remove from a given data sequence. Consequently, since the higher-order moments cannot be reliably estimated from the data, it is important to determine the accuracy of the sample moments and, to this purpose, it should be emphasized that the standard deviations in these estimates depend critically on the shape of the underlying distribution of concentration and, in particular, on the tails

of the distribution. However, the underlying distribution is unknown for concentration fluctuations since this is the quantity that needs to be extracted from the data. In view of this, although the higher-order moments provide the information on the detailed shape of the PDF of concentration, they are not used in this study. Rather, the statistical characterization of the concentration fluctuations are based on the sample quantiles which can be robustly extracted from the input concentration data. The data quantiles of concentration are shown to provide a very convenient, robust and effective method for studying the statistical characteristics of concentration fluctuations and for extracting the shape of the PDF of concentration. Towards this goal, a new model distribution for concentration, namely, the g and h distribution, is introduced and the ability of this distribution to characterize the concentration statistics will be compared with three commonly used model distributions for concentration, namely, the log-normal, the exponential, and the clipped-normal distributions.

DATA DESCRIPTION AND PREPROCESSING

Digital records of concentration fluctuations in ion plumes were supplied to the author by Dr. D. J. Ride. Details of the field experiment from which the concentration data were derived have been described by Ride [9]. The dispersion experiment lasted about 16 minutes and was carried out at Cardington, Bedfordshire, England over flat and unobstructed terrain on 9 July 1985 between 14:00 and 14:30 hours under near neutral stability conditions. The basic data consist of measurements of fluctuations in concentration of a plume of negatively ionized air produced continuously by a relatively powerful ion generator positioned at 1.5 m height. The arrangement of apparatus for the experiment is shown in Figure 1. The concentrations in the ionized plume were measured by four ion collectors placed at a height of 1.5 m and at a distance of 16 m downwind of the ion source. The four ion collectors were positioned in the crosswind direction; collector Nos. 3 and 4 were placed 0.5 m laterally on either side of the line of sight extending from the ion generator along the average wind direction and collector Nos. 1 and 2 were positioned, respectively, at lateral distances of 2.5 and 1.5 m from the line of sight. Consequently, collector Nos. 3 and 4 were positioned approximately at the mean centerline of the ion plume. Ion collector No. 3 malfunctioned and the spurious data obtained from this sensor have not be used in the following analysis. The concentration data were sampled at a rate of 1000 Hz with an analog-to-digital converter with a resolution of 8 bits.

The digital time sequences from collector Nos. 1, 2 and 4 were visually examined and a section consisting of 5 minutes of relatively "clean" concentration data was selected from each sequence for further processing. However, prior to performing the data analysis, some preprocessing of the three data series was necessary. Firstly, the presence of artifacts in the form of narrow spikes of duration 0.001 s or less (essentially single isolated points in the data series which had relatively large values), due probably to the presence of stray electrically charged aerosol particles in the vicinity of the ion collectors, were visually identified and removed from the data sequence. Secondly, small baseline drifts in the data sequences were corrected by

subtracting a pre-selected threshold value from each datum and clipping any negative values to zero. No further filtering was performed on the resulting data sequences and, for the purposes of this study, these preprocessed sequences are considered to consist of ideal, non-time averaged concentration data.

An example of concentration fluctuations in an ion plume is given in Figure 2 which displays 1.5 s (i.e., 1500 data points) of concentration data from collector No. 4 which was positioned approximately at the mean plume centerline. It is clearly seen that the variability in concentration consists of a series of bursts of high concentration interspersed with periods of zero concentration. Note the large spatial gradients in concentration at the leading and trailing edges of the pulses. The intervals of zero concentration illustrate the phenomenon of plume meandering whereby turbulent eddies much larger than the plume move it bodily back and forth over the sampling point resulting in periods when the tracer material is completely absent from the airflow over the sampler. However, eddies comparable in size to the plume redistribute tracer material locally within the plume. An example of in-plume fluctuations is shown in Figure 3 from which it is readily seen that there is considerable small-scale structure in the variation. Observe that even within the instantaneous plume, there is the succession of large concentration or clusters of large concentration separated by sequences of small concentration, the latter consisting of regions of relatively clean air.

A number of standard descriptive statistics were computed for the three data sequences and presented in Table I. These statistics are presented as a function of y/σ_y, where y is the lateral distance from the mean plume centerline and σ_y is the width of the plume; the latter parameter was determined by fitting a Gaussian curve to the mean crosswind concentration profile. Since the data were supplied in terms of digitized voltage levels and since the calibration factors required to convert the voltages to absolute ion concentration levels were not made available to the author, the concentration statistics given in Table I are for data that have been scaled by their respective mean concentration values. The descriptive statistics that have been calculated for the data include the absolute deviation-to mean, the standard deviation-to-mean, the mode-to-mean, the median-to-mean, and the peak-to-mean concentration ratios, as well as the skewness, the kurtosis, and the intermittency γ which is defined as the fraction of the total time that zero concentrations are observed. It should be noted that the kurtosis values used in this paper have had the value 3 subtracted from them—viz., for this "normalized" kurtosis, a value of zero is obtained for a Gaussian distribution. An examination of Table I shows that the concentration data exhibit marked deviations from a Gaussian distribution. In particular, the skewness indicates that the data is strongly skewed to the right and the kurtosis is much larger than in a Gaussian distribution indicating a very peaked (i.e., leptokurtic) distribution. The latter is undoubtedly due to the high intermittency. Along these lines, observe that the mode-to-mean ratio is zero implying that zero concentration is the most frequently occuring value in the data sequence. Furthermore, note that the median-to-mean ratio is zero for all data sequences, indicating that the intermittency is greater than 0.50 for all sequences.

The information on the zero concentration is embodied in the intermittency γ and once this parameter

is known, the intervals of zero concentration represent superfluous information. Consequently, it is convenient to remove the intervals of zero concentration from the unnormalized data sequences and consider the conditional concentration statistics. The latter statistics are presented in Table II. Again, it should be noted that these statistics are calculated for normalized data but, in this case, the data are scaled by the associated conditional means. Observe that the skewness and kurtosis for the conditional concentration data are smaller than the corresponding values for the unconditional concentration data implying that the probability distribution of the conditional data exhibit a smaller departure from a Gaussian distribution. In addition, the variation in the conditional statistics with respect to lateral position y/σ_y is considerably smaller than that in the unconditional statistics. In other words, the statistics of the conditional concentration data scaled by the conditional mean are less sensitive to y/σ_y than the statistics of unconditional concentration data scaled by the unconditional mean. This indicates that the conditional mean constitutes a convenient concentration scale for the conditional data (at least in homogeneous turbulence under neutral atmospheric stability), an observation that is supported by the work of Lewellen and Sykes [7], Sawford [8] and Dinar *et. al.* [10]. Consequently, the following analysis will focus primarily on the normalized conditional concentration data.

SOME EXPLORATORY DATA ANALYSIS

This section focusses on the application of ordered statistics and quantiles to the characterization of the shape of the probability distribution for conditional concentration data. This approach permits a visualization of the distributional shape of the data without the need to consider higher-order moments. As stated previously, moment-based methods for the characterization of distribution shape tend at best to be misleading and, at worst, to be unreliable since these methods are not very robust.

Given the instantaneous conditional concentration data sequence, denoted by $c_1, c_2, c_3, \ldots, c_n$, define a corresponding sequence $c_{(1)}, c_{(2)}, c_{(3)}, \ldots, c_{(n)}$ of the original sequence values permuted to be in ascending numerical order, viz. $c_{(1)} \leq c_{(2)} \leq c_{(3)} \leq \ldots \leq c_{(n)}$. These values are referred to as the order statistics of $c_1, c_2, c_3, \ldots, c_n$. Assume that the conditional concentration data samples are drawn from an underlying population with conditional probability distribution function $F^+(c)$. If p denotes a positive proper fraction, then the quantile of order p (or, equivalently, the p percentile), which will be designated by C_p, is defined as the unique solution of the equation $F^+(C_p) \equiv \Pr(c \leq C_p) = p$. Here, Pr denotes "the probability that". The value of p is referred to as the percentage point. As an example, the quantile of order $1/2$ (or, synonymously, the $1/2$ percentile) is the median of the distribution with $\Pr(c \leq C_{0.5}) = F^+(C_{0.5}) = 1/2$. The quantile function $Q(p)$ ($p \in [0,1]$) corresponding to the distribution function $F^+(c)$ can be now defined as follows:

$$Q(p) = \left(F^+\right)^{-1}(p) \equiv \inf\left\{c : F^+(c) \geq p\right\},$$

where inf denotes infimum. Note that for a fixed value of p, $Q(p) = C_p$. In view of this definition, the quantile function can be estimated from the ordered statistics according to the prescription

$$\hat{Q}(p) = c_{(j)} \qquad \text{for} \qquad \frac{j-1}{n} < p \leq \frac{j}{n}, \quad j = 1, 2, 3, \ldots, n,$$

where the accent $\widehat{(\cdot)}$ denotes the estimate of $(\cdot)$.

A typical example of the sample quantile function for concentration is shown in Figure 4: the sample quantiles for this figure are constructed from the unconditional normalized concentration data from collector No. 4 ($y/\sigma_y = 0.715$). The most distinctive feature of this plot is the long plateau at zero concentration. This long horizontal segment in the sample quantile function suggests the presence of a discrete probability mass at zero concentration. The kink in the quantile function occuring at approximately $p = 0.84$ terminates the plateau region and delineates the beginning of the sharp rise in the quantile function. The location of the kink at $p = 0.84$ coincides numerically with the intermittency γ observed at the sampling point (cf. Table I). The sharp rise in the function after the kink (i.e., "infinite" slope) implies the presence of a zero in the probability density function. Since the probability density function is necessarily non-negative, this suggests that the unconditional concentration data is bimodal. This should not be too surprising since the concentration fluctuations originate from two distinct mechanisms: namely, the plume meandering which gives rise to the intermittency and the within-plume concentration fluctuations. Each of these distinctive mechanisms can be associated with a particular portion of the sample quantile function—the plume intermittency with the long plateau and the within-plume fluctuations with the sharp rise. Thus, the dual structure in the concentration fluctuations is clearly revealed in the sample quantile function.

The sample quantile function for the conditional normalized concentration data obtained from collector No. 4 is displayed in Figure 5. Superimposed on the sample quantile function are two boxes whose widths are equal to the mid-spreads of the lower and upper $1/4$ percentiles (i.e., $\widehat{Q}(0.25)$ and $\widehat{Q}(0.75)$) and the $1/8$ percentiles (i.e., $\widehat{Q}(0.125)$ and $\widehat{Q}(0.875)$). The mid-spread is defined as the upper percentile value minus the corresponding lower percentile value. The solid horizontal line ruled inside the smaller box delineates the $1/2$ percentile or median for the data. The display in Figure 5 is a quantile box plot for the concentration data. The box plot was first applied to statistical data analysis by Tukey [11]. Observe that the sample quantile function for the conditional concentration data is smooth and does not exhibit the long horizontal plateau that was so distinctive in the quantile function for the unconditional concentration data (cf. Figure 4). The shape of the sample quantile function for the conditional concentration data is quite symmetrical within the two quantile boxes (cf. Figure 5). However, outside these two quantile boxes, the asymmetry in the data is readily evident. Indeed, note that the data which comprise the right tail increase more rapidly than the data which comprise the left tail, viz. the right tail is longer than the left. The latter fact, in itself, should not be too surprising since the concentration cannot take on negative values.

The density-quantile function $fQ(p)$ defined as

$$fQ(p) \equiv p^+(c = Q(p)),$$

where $p^+(c)$ denotes the probability density function for the conditional concentration data, can be estimated

from the sample quantile function as follows:

$$\hat{f}Q(k\Delta) = \frac{2\Delta}{\hat{Q}((k+1)\Delta) - \hat{Q}((k-1)\Delta)}, \tag{1}$$

where Δ is a constant step-size and the density-quantile function is evaluated at a sequence of equi-spaced percentage points denoted $p = \Delta, 2\Delta, 3\Delta, \ldots, 1-2\Delta, 1-\Delta$. Note that the density-quantile function is simply the probability density function evaluated at the quantiles. Figure 6 shows the sample density-quantile function for the normalized conditional concentration data from collector No. 4. This plot was constructed by estimating points on the density-quantile function using Equation (1) and the results contained in Figure 5. The sample density-quantile in Figure 6 was plotted by interpolating between the estimated points using a cubic spline interpolation routine. Observe that the density-quantile function is unimodal; the maximum value of the density-quantile function occurs at approximately $c/C = 0.06$ which is in accordance with the mode-to-mean ratio of 0.058 calculated earlier (cf. Table II). Furthermore, the mode of the density-quantile function is very sharp and narrow. The data is clearly skewed to the right and, as has been observed earlier in conjunction with the quantile function, the right tail of the distribution is somewhat longer than the left tail.

To investigate further the shape of the distribution of concentration, certain selected quantiles were extracted from the ordered statistics. The selected quantiles consisted of the $1/2$ percentile ($p = 0.5$), the lower and upper $1/4$ percentiles ($p = 0.25$ and $p = 0.75$), the lower and upper $1/8$ percentiles ($p = 0.125$ and $p = 0.875$), the lower and upper $1/16$ percentiles ($p = 0.0625$ and $p = 0.9375$) and, so forth, continuing outward into the tails by successively halving the tail area on each end. The following quantitative diagnostic measures were computed from these selected quantiles: the mid-summaries $\mu(p)$ defined as

$$\mu(p) = \frac{1}{2}\Big(Q(1-p) + Q(p)\Big), \qquad 0 \le p \le 0.5;$$

the mid-spreads defined as

$$\delta(p) = Q(1-p) - Q(p), \qquad 0 \le p \le 0.5;$$

and the pseudosigmas defined as

$$\sigma(p) = \delta(p)/(z_{1-p} - z_p),$$

where z_p denotes the p-th order quantile of the standard Gaussian distribution (i.e., a Gaussian distribution with mean zero and unit standard deviation). It should be noted that $\sigma(p)$ can be interpreted as the ratio of the mid-spread for the data to the corresponding mid-spread for the standard Gaussian distribution and, as such, is a measure of the elongation of the tails of the data when compared to the Gaussian standard. Taken together, the mid-summaries and pseudosigmas are useful measures for describing two important characteristics of the distribution shape of the data: namely, skewness (departure from symmetry) and elongation (weight in the tails).

Figure 7 shows the selected mid-summaries $\mu(p)$ for the normalized conditional concentration data from collector No. 4 plotted against the square of the corresponding standard Gaussian quantile z_p^2. If the underlying distribution for the concentration data is symmetric, then the upper and lower p-th order quantiles should be symmetrically located with respect to the median. Consequently, for a symmetric distribution, the mid-summaries $\mu(p)$ for each value of p should equal a constant, that constant being the median. Hence, in view of this, if the concentration data were perfectly symmetric, the mid-summaries versus z^2 plot (or mid-versus-z^2 plot for short) displayed in Figure 7 would be a horizontal line. However, an examination of Figure 7 shows that the mid-versus-z^2 plot of the concentration data increases as one moves farther into the tails implying that the data are skewed to the right. Furthermore, note that points on the mid-versus-z^2 plot seem to determine two straight-line segments—the first seven points (embodying information up to the lower and upper $1/128$ percentiles) fall on a line of large positive slope (indicating a substantial skewness to the right) and the last six points (embodying information from the lower and upper $1/128$ percentiles to the lower and upper $1/8192$ percentiles) fall on a line of small positive slope (indicating only a slight skewness to the right). Hence, concentration data associated with the extreme lower and upper tails of the distribution exhibit only a minor departure from symmetry. The latter is probably due to the fact that concentration is bounded below by zero (negative concentrations cannot occur) and above by the initial source concentration (maximum concentration corresponding to undiluted material). Consequently, the mid-versus-z^2 plot (cf. Figure 7) shows that the skewness in the concentration data is confined primarily to the near tails of the distribution.

Figure 8 displays a plot of the selected pseudosigmas $\sigma(p)$ versus z_p^2 for the normalized conditional concentration data from collector No. 4. This plot can be used to examine the tailweight of the distribution of concentration. A horizontal line on the pseudosigma-versus-z^2 plot would indicate that the data are neutrally elongated in comparison to the standard Gaussian distribution. A perusal of Figure 8 shows that $\sigma(p)$ decreases systematically implying that the distribution for concentration is less elongated (i.e., has a lighter or shorter tail) than the standard Gaussian distribution.

Unlike skewness, elongation can be characterized by studying the upper and lower halves of the data separately; indeed, it is not necessary that the behavior of the lower tail of a distribution be identical to that of the upper tail [11]. To this end, consider a pushback analysis whereby the following "flattened" data quantile values (i.e. pushback values) f_p are computed:

$$f_p = Q(p) - sz_p, \qquad 0 \le p \le 1,$$

where $Q(p)$ is the p-th order data quantile, z_p is the p-th order standard Gaussian quantile and s is a scale for the data which is estimated using

$$s = \text{med}\Big\{ \big(Q(1-p) - Q(p)\big) / \big(z_{1-p} - z_p\big) \Big\}, \qquad 0 \le p \le 0.5.$$

Here, med denotes median. Observe that s is simply the median of the pseudosigmas $\sigma(p)$ of the data. Now, it should be noted that a plot of f_p versus z_p will yield a horizontal line if the concentration data is Gaussian with scale s. Furthermore, if the concentration data is Gaussian with a scale different than s, then the f_p versus z_p plot will result in a straight line with non-zero slope. The slope of this line is directly related to the scale of the Gaussian.

Figure 9 shows the pushback normalized conditional concentration values (collector No. 4) for selected p-th order quantiles plotted as a function of z_p. The scale $s = 0.74$ for the concentration data was obtained by selecting the median of the pseudosigma values displayed in Figure 8. A careful examination of Figure 9 indicates that the points fall on three distinct straight-line segments. The first (i.e., leftmost) twelve points determine a line with slope 0.79 and correspond to the information in the lower tail (i.e., below the median) of the distribution. The next seven points (which includes the median point) fall on a line with slope 0.47 and determine the information in the near upper tail of the distribution. The last (i.e., rightmost) six points are shifted downwards and define a line with slope -0.31; these points embody the information in the extreme upper tail of the distribution. Consequently, the distribution of concentration can be characterized by three Gaussians with different scales (as evidenced by the three straight lines of different slopes in Figure 9). From the slopes of these lines, it can be determined that the data in the lower tail is Gaussian with a scale (in c/C units) that is approximately $(s + 0.79)/(s + 0.47) = 1.26$ times the scale of the near upper tail; similarly, the data in the extreme upper tail is Gaussian with a scale (in c/C units) that is about $(s - 0.31)/(s + 0.47) = 0.355$ times the scale of the near upper tail. Since the scale in the lower tail is only 26 percent larger than that in the near upper tail, it follows that the concentration distribution can be well approximated up to about the 127/128 percentile with a single truncated Gaussian distribution. However, this single Gaussian approximation will not be able to properly characterize the extreme upper tail of the concentration distribution. Finally, it is useful to point out that the behavior of the upper tail of the concentration distribution as displayed in Figure 9 is consistent with the information concerning the tail as displayed in Figure 7. In particular, note that the breakpoint that separates the near upper tail from the extreme upper tail occurs at the same point in both figures.

MODELLING CONCENTRATION DISTRIBUTIONS

It is known that the dependence of the conditional probability distribution on the crosswind distance from the mean plume centerline can be removed by scaling the conditional concentration data c by the conditional mean $\bar{c}$ (Lewellen and Sykes [7]; Sawford [8]; Dinar *et. al.* [10]). Figure 10 displays the conditional probability distributions of the normalized concentration data obtained from the three ion collectors. This figure shows that the three probability distributions virtually coincide. However, this coincidence is somewhat misleading since the plot of Figure 10 is more sensitive to differences in the probability distributions in the center or middle of the distributions than in the tails. Consequently, to study the differences in the

probability distributions in the tails, it is necessary to use another form of display. To this purpose, consider a quantile-quantile (Q-Q) plot which is simply a plot of the quantiles of one distribution against those of another. Since the quantile is a rapidly varying function of the percentage point p where the density is sparse, the Q-Q plot is particularly effective in delineating the differences in the tails of two distributions. Figure 11 exhibits the Q-Q plot of the sample quantiles for the normalized conditional concentration data from collector Nos. 1 ($y/\sigma_y = 3.575$) and 2 ($y/\sigma_y = 2.145$) against those from collector No. 4 ($y/\sigma_y = 0.715$). This figure clearly shows that probability distributions for c/C are not completely independent of y/σ_y; indeed, the variation of the behavior of these probability distributions with crosswind distance becomes more acute at the extreme upper tails.

The two basic mechanisms responsible for concentration fluctuations, namely, plume meandering resulting in the intermittent periods of zero concentration and in-plume mixing and interleaving of contaminant and clean air resulting in fine-scale internal variations, are mutually exclusive, so, the probability distribution of concentration at a fixed receptor point can be written as

$$\begin{aligned} F(c) &= \gamma F^-(c) + (1-\gamma)F^+(c) \\ &= \gamma H(c) + (1-\gamma)F^+(c). \end{aligned} \tag{2}$$

Here $F(c)$, $F^-(c)$ and $F^+(c)$ are, respectively, the probability distribution for the total concentration c, for the gaps of zero concentration, and for the in-plume concentration fluctuations. Since there is a finite probability for observing periods of zero concentration, the probability distribution $F^-(c)$ in Equation (2) is expressed as a unit step function $H(c)$ with a step at $c = 0$ of magnitude γ. In view of Equation (2), the probability density function can be represented here by

$$\begin{aligned} p(c) \equiv \frac{dF(c)}{dc} &= \gamma p^-(c) + (1-\gamma)p^+(c) \\ &= \gamma\delta(c) + (1-\gamma)p^+(c), \end{aligned} \tag{3}$$

where $p(c)$ is the total PDF of concentration, $p^-(c)$ is the PDF for zero concentration, and $p^+(c)$ is the PDF for in-plume fluctuations (viz., the conditional PDF for concentration fluctuations with the intervals of zero concentration censored). Observe that $p^-(c) = \delta(c)$, the Dirac delta function, which results from the step discontinuity (i.e. $H(c)$) in $F(c)$.

A number of models have been proposed for the conditional probability distribution, the most important of which are the log-normal ([3], [4]), the exponential ([5], [6]) and the clipped-normal ([7], [8], [9], and [10]) distributions. In addition to these common distributions, this paper focusses on the g and h distribution. The g and h family of distributions, obtained by a nonlinear transformation of a standard Gaussian process, is extremely flexible in characterizing a great variety of distribution shapes [12,13]. The probability distribution $F^+(c)$ or, equivalently, the probability density function $p^+(c)$, follows the g and h distribution

if the random variable c is obtained by transforming the standard normal variable z according to

$$c = m + s\frac{\exp(gz) - 1}{g}\exp\left(hz^2/2\right), \tag{4}$$

where g and h are functions of z^2 and m and s are location and scale parameters for c. In this context, it should be noted that m is the median of c. The function g controls the asymmetry or skewness of the distribution and, for the purposes of the present paper, g will be taken to be the following simple polynomial in z^2:

$$g = g_0 + g_1 z^2, \tag{5}$$

where g_0 and g_1 are constants. The function h controls the amount of elongation or the amount of tail-stretching in the distribution. In the present paper, the function h will be chosen as the following simple polynomial in z^2:

$$h = h_0 + h_1 z^2, \tag{6}$$

where h_0 and h_1 are constants. Other more general forms for g and h as functions of z^2 are possible and a discussion on the latter point can be found in Tukey [12] and Hoaglin and Peters [13]. Note that when $h_0 = 0$, $h_1 = 0$ and $g_1 = 0$, the g and h distribution reduces to the three-parameter log-normal distribution provided g_0 is positive, the latter constraint ensuring the positive skewness (i.e., skewness to the right) of the distribution.

The four probability distribution models were fitted to the conditional concentration (i.e., c/C) data for collector No. 4 ($y/\sigma_y = 0.715$). The parameters of the log-normal, the exponential, and the clipped-normal distributions were obtained from the first- and/or second-order moments of the concentration data. The parameters of the g and h distribution were computed by matching a number of theoretical quantiles for the distribution with the corresponding sample quantiles extracted from the data as described by Hoaglin and Peters [13]. The results of this analysis are summarized in Table III. The parameters displayed in this table include $\bar{c}$ (i.e. C), the mean of the exponential distribution; μ_g and σ_g, the logarithmic mean and standard deviation of the log-normal distribution; and μ_c and σ_c, the clipped mean and standard deviation of the clipped-normal distribution. Table IV compares selected quantiles of the fitted distributions with the sample quantiles computed from the data. Observe that the g and h distribution provide an accurate characterization of the data quantiles with a relative error of less than 2 percent.

The results of Table IV can be graphically displayed in the form of Q-Q plots of the data quantiles versus the fitted quantiles of the model distributions. On such a plot, it should be noted that an exact fit between the concentration distribution and the model distribution results in a straight line which has slope 1 and intercept 0 when the ordinate (data quantiles) and abscissa (fitted quantiles) have equal scales. Figure 12 illustrates the Q-Q plot for the concentration data and the fitted log-normal distribution. The extreme curvature of the Q-Q plot in Figure 12 implies that the log-normal distribution is a poor fit to

the concentration data; indeed, the log-normal distribution grossly overestimates the data quantiles in the extreme upper tails. Consequently, while the data exhibits periods of high concentration, the fraction of extreme values at these high concentrations is not so large as to support the application of a log-normal distribution. Figure 13 shows the Q-Q plot of the concentration data quantiles with the fitted exponential, clipped-normal and g and h distributions chosen as the reference distributions. An examination of Figure 13 shows that the g and h distribution is superior to both the exponential and clipped-normal distributions in modelling the concentration statistics. Indeed, the linearity of the Q-Q plot configuration for the sample quantiles versus those for the g and h distribution provides the evidence that this distribution is a good probability model for the concentration data. Along these lines, it should be noted that a least squares fit of the points gives a slope of 1.007 ± 0.006 and an intercept -0.013 ± 0.019 with a correlation coefficient r between the data and fitted quantiles of 0.999. The concave r-shaped curves of the Q-Q plots of the data quantiles versus the quantiles from the exponential and clipped-normal distributions show that these model distributions have longer or heavier tails than the actual underlying distribution of the concentration data. However, the clipped-normal distribution provides a better approximation for the data than the exponential distribution. Observe that the clipped-normal distribution provides a reasonable description for the concentration data in the lower and near upper tails. This fact has already been indicated with regard to the pushback analysis performed earlier and summarized in Figure 9. Furthermore, in adherence with Figure 9, the clipped-normal distribution does not provide an accurate description of the concentration data in the extreme upper tail. However, it should be remarked that the clipped-normal distribution does provide a reasonably accurate and relatively conservative model for the concentration data (i.e., it overpredicts the concentration in the extreme upper tail). As an example, it overpredicts the 0.996 percentile of the concentration by about 15 percent, although the overprediction becomes progressively worse as one proceeds outward into the extreme upper tail.

Figures 14, 15 and 16 show the unconditional frequency distribution data superimposed on the total model probability distribution obtained from Equation (2) using the observed intermittency γ and the fitted conditional clipped-normal and g and h distributions at, respectively, $y/\sigma_y = 0.715$, 2.145, and 3.575. Both the clipped-normal and g and h distributions provide good approximations for the concentration data, although the g and h distribution is clearly seen to model to data more accurately in the upper tail. Again, it is important to note that the plots in these figures tend to emphasize the appearance of differences in the centers of the distributions. Consequently, these plots are somewhat misleading in that they seem to indicate that both the clipped-normal and g and h distributions provide equally good approximations for the data at the upper tails. However, a perusal of the Q-Q plot of Figure 13 clearly shows that this is not so. Indeed, Figure 13 illustrates that the theoretical quantiles of the clipped-normal distribution bends toward the horizontal at the upper end suggesting that the concentration data is negatively elongated relative to the clipped-normal. In other words, the curve flattens away from the ideal straight line at the upper end

implying that the concentration data quantiles are less extreme than those predicted by the clipped-normal distribution.

Figures 17 and 18 show the conditional model probability density functions (clipped-normal and g and h distribution) obtained from fitting the scaled concentration data at approximately the plume centerline ($y/\sigma_y = 0.715$) and the plume edge ($y/\sigma_y = 3.575$). The clipped-normal PDF is remarkably similiar in form to the g and h PDF with the exception of the presence of a single sharp maximum or peak in the latter that is absent in the former. The clipped-normal PDF for the conditional normalized concentration data attains its maximum at $c/C = 0$. On the other hand, the g and h PDF has a maximum at approximately $c/C = 0.60$ for $y/\sigma_y = 0.715$ and at $c/C = 0.12$ for $y/\sigma_y = 3.575$. The maximum attained by the g and h PDFs at the plume centerline and edge agree well with the ratio of the mode-to-mean values of 0.058 and 0.115 obtained from the computation of the descriptive conditional concentration statistics (cf. Table II). Observe that the shape of the conditional PDF of concentration at $y/\sigma = 0.715$ is similar to that at $y/\sigma_y = 3.575$, suggesting that the fine-scale structures of the fluctuations across the plume are reasonably self-similar. Furthermore, the shape of the conditional PDF given by the g and h distribution is in agreement with the shape of the empirical density-quantile function illustrated in Figure 6 which clearly shows the existence of a sharp and narrow peak at $c/C \neq 0$; in this respect, the clipped-normal distribution is not consistent with shape of the empirical density-quantile function.

When the conditional clipped-normal and g and h PDFs are inserted into Equation (3), it is clear that the resulting total PDF $p(c)$ is unimodal for the clipped-normal distribution model and bimodal for the g and h distribution model. It is important to state that the bimodal model for $p(c)$ is consistent with the empirical quantile function depicted in Figure 3. In this regard, the unimodal model for $p(c)$ does not account for the sharp rise in the graph of the sample quantile function shown in Figure 3. Furthermore, the bimodal form of the total PDF $p(c)$, predicted by the g and h distribution, agrees with the bimodality of the PDF as predicted by the physically motivated strand model developed by Chatwin and Sullivan [14] in the context of clouds. However, this model can be extended to the case of plumes by simple analogy and reinterpretation of certain results. In this model, the contaminant material injected into a turbulent flow field is visualized to be drawn out into a complex entanglement of highly distorted sheets and strands which are extended in one direction and compressed in another. The thinning of these contorted sheets and strands result in sharply varying values of concentration in the contaminant cloud—the sharp gradients in concentration serve to enhance the effects of molecular diffusion and this process continues until the sheets and strands have a minimum thickness comparable to the conduction cut-off length, at which point the thinning due to stretching is balanced by the smearing and smoothing out of the sharp interfaces between the filaments and the uncontaminated fluid due to molecular diffusion. Chatwin and Sullivan [14] use similarity arguments to show that sampling such a cloud with a probe of volume a^3 results in a bimodal PDF—the first mode at $c = 0$ corresponds to the sampling of uncontaminated fluid and the second mode at $c = c(a)$ delineates

the sampling of contaminant material within the strands. The PDF $p(c)$ is shown essentially to decrease monotonically as $c \rightarrow c(0)$ $(c \geq c(a))$ where $c(0)$ is the (maximum) concentration at the strand centerline. In this regard, it should be noted that the shape of the PDF predicted by the strand model is consistent with the shape of the concentration PDF modelled by the g and h distribution (cf. Figures 17 and 18 which corresponds to the "toe" of the total PDF $p(c)$ where $c \geq c(a)$).

SUMMARY AND CONCLUSIONS

In this paper, the shape of the probability distribution of concentration fluctuations is studied empirically using high resolution ion plume dispersion data. To characterize the physical nature of the processes responsible for these fluctuations, it is necessary and significant to measure or calculate higher-order moments of the process since these moments embody the information from the tails of the probability distribution. However, it is difficult to estimate these moments reliably from concentration fluctuation data since they are strongly affected by outliers in the data. This realization, coupled with the fact that the precise shape of the probability distribution in the tails seem to require knowledge of the higher-order moments with inordinately high degrees of precision, has resulted in the consideration, instead, of the order statistics as a vehicle for the useful description and characterization of the distribution shape. Unlike the higher-order moments, the order statistics for the concentration data can be determined robustly from even a moderately-size batch of data. In this regard, it is shown that the order statistics and the selection of various quantiles from these statistics form the basis for a range of useful graphical methods for the exploratory and confirmatory data analysis aimed toward the study of the distributional shape of concentration data and the distributional relationships between the data and the model (theoretical) frequency distributions used to describe them. Indeed, the order statistics of the concentration data provide an exhaustive representation of the data that not only lends itself to graphical representation and visualization, but also provides a robust characterization of location, scale, and shape of the data. In this respect, the order statistics provide a valuable tool for studying concentration fluctuation data with the viewpoint of assessing the symmetry and tailweight in the data and for revealing peculiarities in the data.

The data analysis techniques based on ordered statistics have been applied to some high resolution (up to 1000 Hz bandwidth) concentration fluctuation measurements obtained with an ion tracer. A sample quantile function constructed from the concentration data clearly shows the bimodality of the underlying probability distribution for the process. The probability distribution of the conditional normalized concentration data (i.e., data with the zero concentration censored and scaled by the conditional mean) is relatively insensitive to the lateral location in the plume. However, this invariance of distribution shape with lateral position applies primarily to the center and shoulders of the distribution. Indeed, a Q-Q plot of the conditional concentration data quantiles at one lateral position versus those at another clearly reveals the deviations in the distribution shape in the upper tails. A knowledge of the dependence of the shape of the

probability distribution on location in the plume is important, since the accurate assessment of the risks and hazards caused by the accidental or deliberate release of certain highly toxic or flammable gases requires the accurate prediction of the PDF of concentration, especially in the upper tails. A pushback analysis of the conditional concentration data shows that the data can be well approximated by a single (truncated) Gaussian distribution in the lower and near upper tails. However, this single Gaussian cannot accurately represent the important information contained in the extreme upper tails.

Several probability distributions have been considered as models for the concentration fluctuation data: these include three commonly used models, namely, the log-normal, the exponential, and the clipped-normal distributions and a new model, namely, the g and h distribution. The log-normal distribution is not consistent with the concentration data and results in a considerable overestimation of the frequency of occurrence of high concentrations. The exponential distribution also overestimates the frequency of occurence at the upper end of the distribution, although this overestimate is certainly not as severe as that provided by the log-normal distribution. The clipped-normal distribution provides a reasonably accurate representation of the concentration data in the lower and near upper ends of the distribution, but overestimates the frequency of high concentration in the extreme upper end. However, this overestimation is certainly tolerable and, in this respect, the clipped-normal constitutes a useful and moderately conservative model for concentration fluctuations. The g and h distribution, based on the nonlinear transformation of a Gaussian process, is clearly superior to the more common log-normal, exponential, and clipped-normal distributions as a model for concentration fluctuations. The g and h distribution is able to accurately represent the nature of the variations in concentration in the both the lower and upper tails. Indeed, the g and h distribution provides a convenient and effective model for studying the statistical characteristics of concentration fluctuations and for accurately characterizing the random process under consideration. Furthermore, the shape of the g and h distribution is consistent with the shape of the PDF predicted using the strand model developed by Chatwin and Sullivan [14].

The g and h distribution provides an accurate representation for the conditional concentration statistics. However, it is not possible to specify the g and h distribution using only the information embodied in the first- and second-order moments of the concentration fluctuations. In view of this, it is not sufficient to model only the mean concentration, the relative fluctuation intensity, and the intermittency in order to specify the total probability distribution of concentration. A number of models (e.g. Wilson *et. al.* [15], Durbin [16], and Sawford [17]) have been formulated for the prediction of concentration fluctuation statistics, but, these models have been restricted to the prediction of the first- and second-order moments only. As noted before, a realistic model of the random processes governing concentration variations require the prediction of higher-order moments, or, better yet, the probability distribution itself. In this regard, it should be noted that the Lagrangian model recently proposed by Kaplan and Dinar [18] attempts to fulfil the latter objective, although these results only have been considered in the context of one-dimensional homogeneous turbulence.

REFERENCES

1. Pasquill, F., Atmospheric Diffusion: The Dispersion of Windborne Material from Industrial and other Sources (2nd Edition), Chichester: Ellis Horwood Limited, 1974.

2. Chatwin, P. C.,"The Use of Statistics in Describing and Predicting the Effects of Dispersing Gas Clouds", J. Haz. Materials, Vol. 6, pp. 213–230, 1982.

3. Csanady, G. T., Turbulent Diffusion in the Environment, Dordrecht, Holland: D. Reidel Publ. Co., 1973.

4. Jones, C. D., "On the Structure of Instantaneous Plumes in the Atmosphere", J. Haz. Materials, Vol. 7, pp. 87–112, 1983.

5. Barry, P. J., "Stochastic Properties of Atmospheric Diffusivity", AECL-5012, Atomic Energy of Canada Ltd., Chalk River, Ontario, Canada, 1975.

6. Hanna, S. R., "The Exponential Probability Density Function and Concentration Fluctuations in Smoke Plumes", Boundary-Layer Meteorol., Vol. 37, pp. 89–106, 1986.

7. Lewellen, W. S. and Sykes, R. I., "Analysis of Concentration Fluctuations for Lidar Observations of Atmospheric Plumes", J. Clim. Appl. Meteorol., Vol. 25, pp. 1145–1154, 1986.

8. Sawford, B. L., "Conditional Concentration Statistics for Surface Plumes in the Atmospheric Boundary Layer", Boundary-Layer Meteorol., Vol. 38, pp. 209–223, 1987.

9. Ride, D. J., "Modelling Fluctuations in the Concentration of Neutrally Bouyant Substances in the Atmosphere", Ph.D. thesis, University of Liverpool, 1987.

10. Dinar, N., Kaplan, H. and Kleiman, M., "Characterization of Concentration Fluctuations of a Surface Plume in a Neutral Boundary Layer", Boundary-Layer Meteorol., Vol. 45, pp. 157–175, 1988.

11. Tukey, J. W., Exploratory Data Analysis, Reading, Mass.: Addison-Wesley Publishing Co., 1977.

12. Tukey, J. W., Modern Techniques in Data Analysis, NSF-sponsored regional research conference at Southeastern Massachusetts University, North Dartmouth, Mass., 1977.

13. Hoaglin, D. C. and Peters, S. C., "Software for Exploring Distribution Shape" in Proceedings of Computer Science and Statistics: Twelfth Annual Symposium on the Interface (Gentleman, J. F., ed.), University of Waterloo, Ontario, pp. 418–423, 1979.

14. Chatwin, P. C. and Sullivan P. J., "On the Probability-Density Function of Concentration in Turbulent Diffusion", Transactions of the CSME, Vol. 5, pp. 192-196, 1979.

15. Wilson, D. J., Robins, A. G. and Fackrell, J. E., "Intermittency and Conditionally-Averaged Concentration Fluctuation Statistics in Plumes", Atmos. Environ., Vol. 17, pp. 1053-1064, 1985.

16. Durbin, P. A., "A Stochastic Model of Two-Particle Dispersion and Concentration Fluctuations in Homogeneous Turbulence", J. Fluid Mech., Vol. 100, pp. 279-302, 1980.

17. Sawford, B. L., "Lagrangian Statistical Simulation of Concentration Mean and Fluctuation Fields", J. Climate Appl. Met., Vol. 24, pp. 1152-1166, 1985.

18. Kaplan, H. and Dinar, N., "A Stochastic Model for Dispersion and Concentration Distribution in Homogeneous Turbulence", J. Fluid Mech., Vol. 190, pp. 121-140, 1988.

TABLE I

Some descriptive unconditional concentration statistics.

Statistic	Sensor 1 ($y/\sigma_y = 3.575$)	Sensor 2 ($y/\sigma_y = 2.145$)	Sensor 4 ($y/\sigma_y = 0.715$)
(Abs. Deviation)/Mean	1.975	1.940	1.700
(Std. Deviation)/Mean	12.53	7.985	3.160
Mode/Mean	0.000	0.000	0.000
Median/Mean	0.000	0.000	0.000
Peak/Mean	324.5	194.1	27.42
Skewness	16.79	11.87	3.916
Kurtosis	320.1	180.3	16.61
Intermittency	0.988	0.972	0.838

TABLE II

Some descriptive conditional concentration statistics.

Statistic	Sensor 1 ($y/\sigma_y = 3.575$)	Sensor 2 ($y/\sigma_y = 2.145$)	Sensor 4 ($y/\sigma_y = 0.715$)
(Abs. Deviation)/Mean	0.710	0.639	0.717
(Std. Deviation)/Mean	0.939	0.892	0.884
Mode/Mean	0.115	0.085	0.058
Median/Mean	0.670	0.780	0.751
Peak/Mean	3.869	5.381	4.448
Skewness	1.351	1.979	1.061
Kurtosis	0.807	0.786	0.479

TABLE III

Parameters for some model probability distributions for conditional concentration data (Sensor 4).

Distribution	Parameter values
Exponential	$\hat{c}\,(C) = 17.3092$
Log-normal	$\mu_g = 2.3387$
	$\sigma_g = 1.1482$
Clipped-normal	$\mu_c = -3.5654$
	$\sigma_c = 2.3123$
g and h	$m = 13.0000$
	$s = 16.0286$
	$g_0 = 0.7326$
	$g_1 = -0.0215$
	$h_0 = -0.2367$
	$h_1 = 0.0080$

TABLE IV

Comparison of data quantiles (Sensor 4) with fitted quantiles
for log-normal, clipped-normal, and g and h distributions.

p	Data quantile (c/C)	Log-normal	Clipped-normal	g and h
0.0625	0.0578	0.1029	0.0750	0.0695
0.1250	0.1155	0.1599	0.1541	0.1224
0.2500	0.2889	0.2761	0.3269	0.2822
0.5000	0.7510	0.5990	0.7575	0.7510
0.7500	1.5600	1.2990	1.4291	1.5140
0.8750	2.1376	2.2440	2.0376	2.1639
0.9375	2.6000	3.4866	2.5974	2.6910
0.96875	3.0620	5.0847	3.1179	3.1063
0.984375	3.4086	7.1036	3.6066	3.4264
0.9921875	3.6974	9.6156	4.0684	3.6689
0.9960938	3.9285	12.703	4.5077	3.8506
0.9980469	3.9863	16.457	4.9211	3.9863
0.9990235	4.1019	20.986	5.3227	4.0891
0.9995118	4.1885	26.400	5.6976	4.1699
0.9997559	4.3041	32.832	6.0791	4.2379
0.999878	4.3330	40.426	6.4069	4.3009

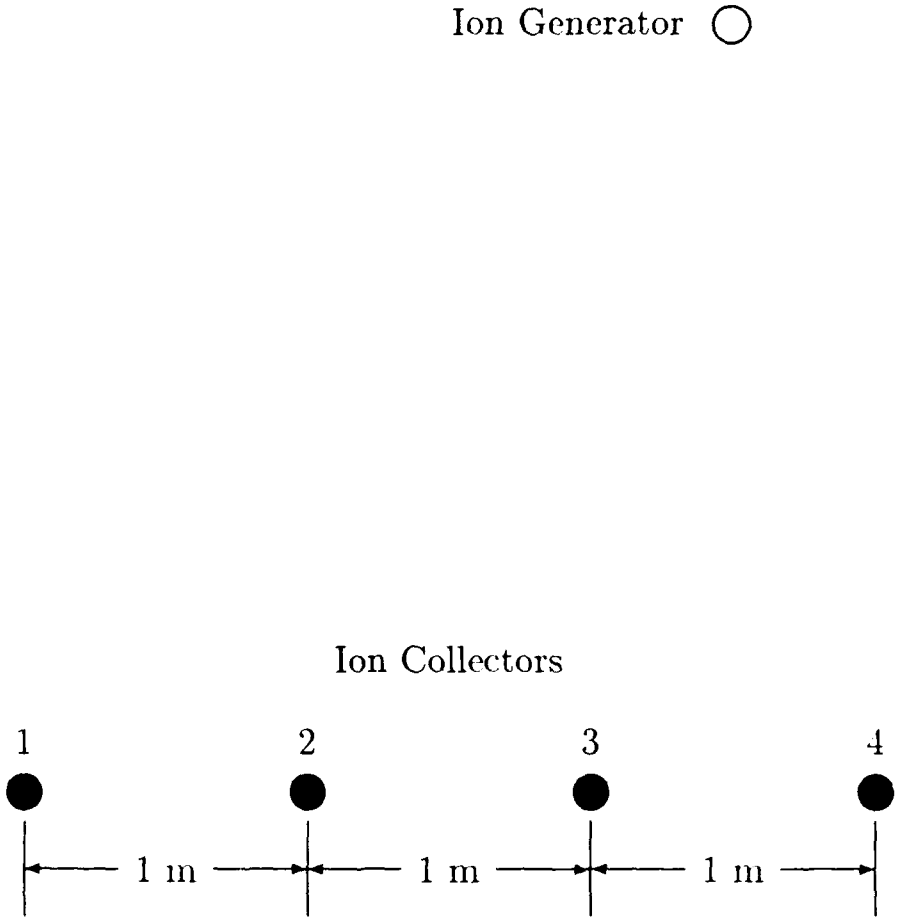

FIGURE 1

Schematic of experimental setup. Mean wind direction was from the ion generator and directed along a ray perpendicular to the line formed by the array of 4 ion collectors.

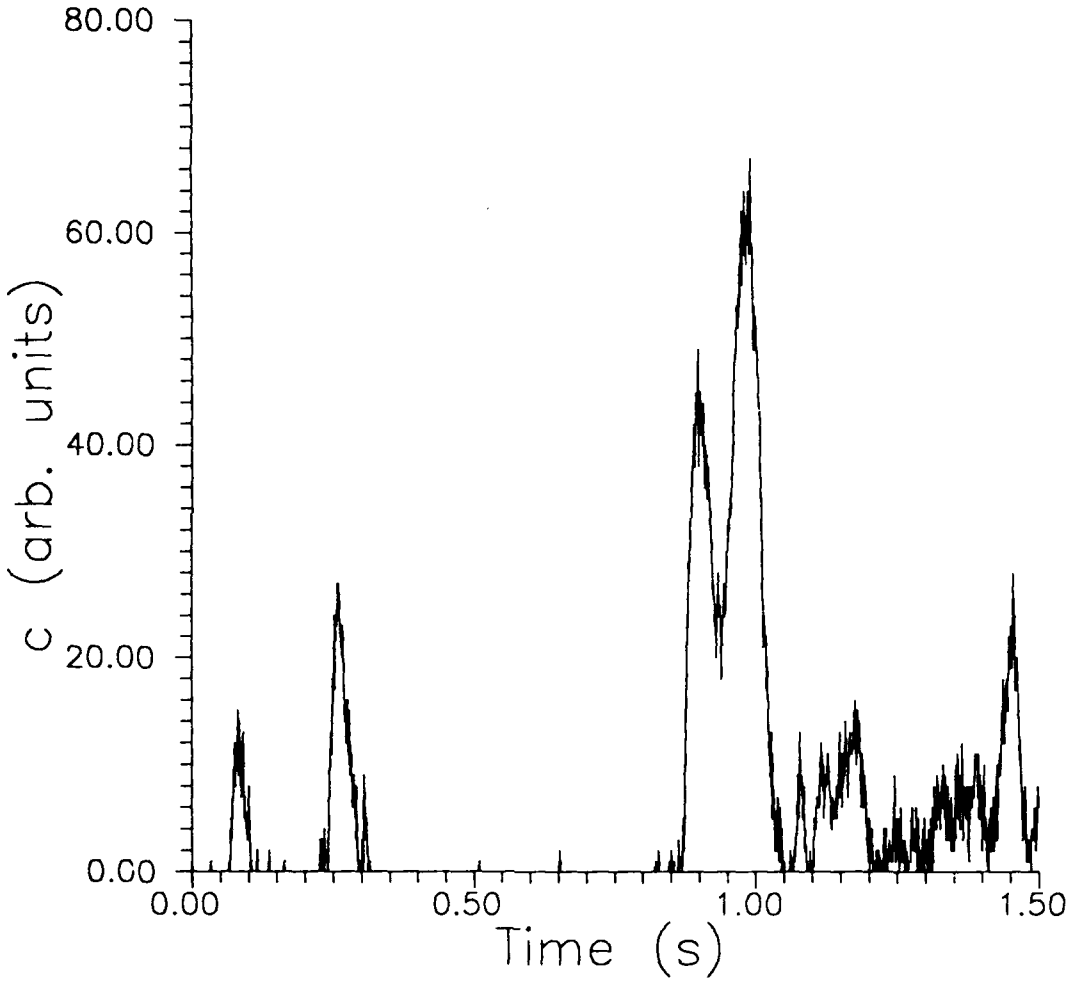

FIGURE 2

An example of a concentration time series illustrating plume intermittency (associated with zero concentration periods) and in-plume concentration fluctuations (corresponding to the presence of short bursts of large concentration).

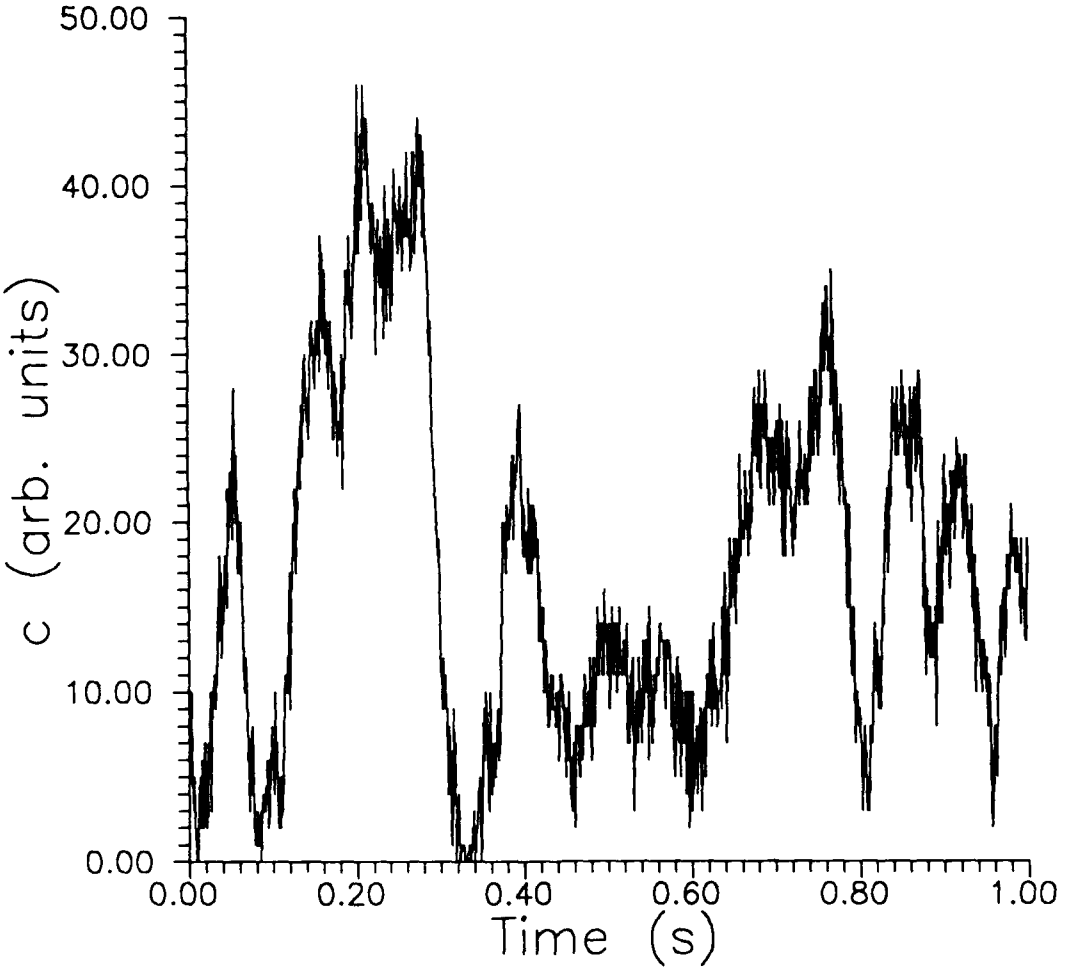

FIGURE 3

An example of a segment of a concentration record illustrating in-plume fine-scale turbulent fluctuations.

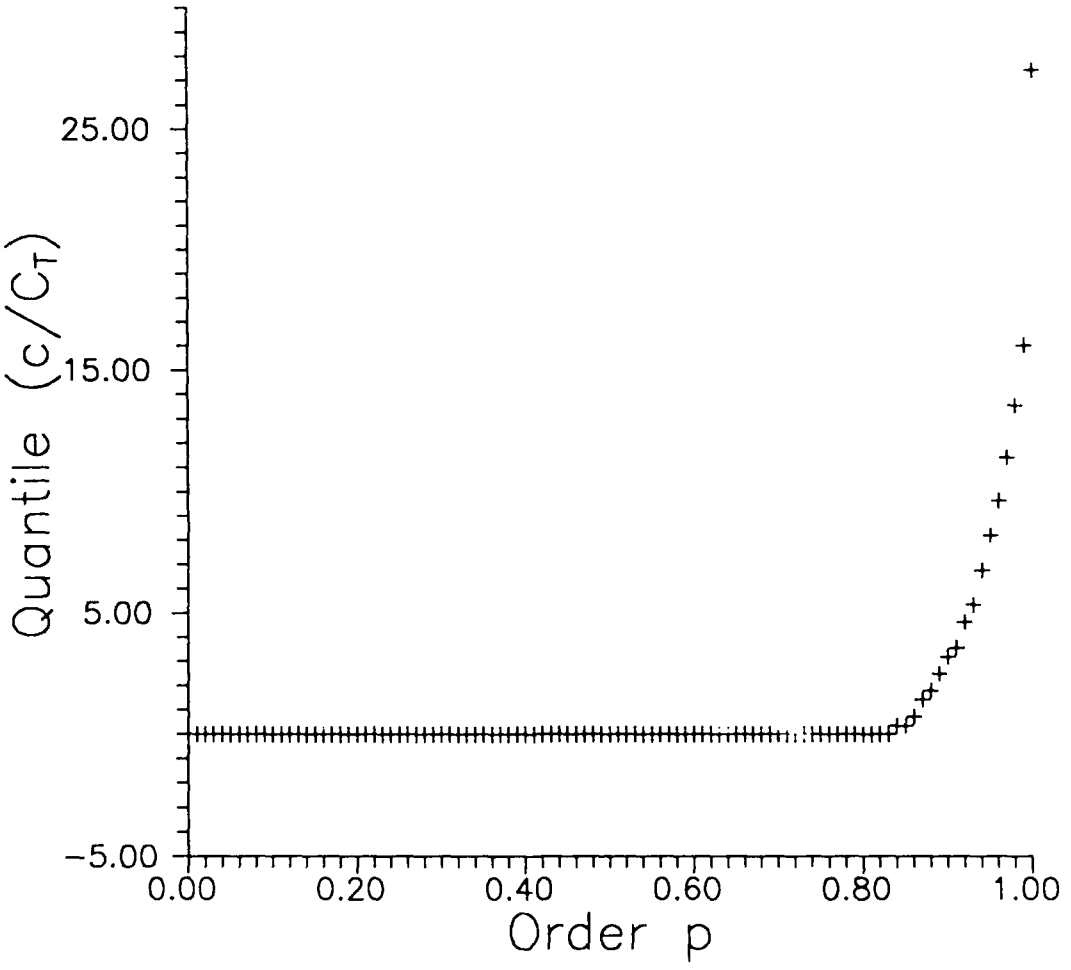

FIGURE 4

The sample quantile function for the unconditional normalized concentration time series from Sensor 4.

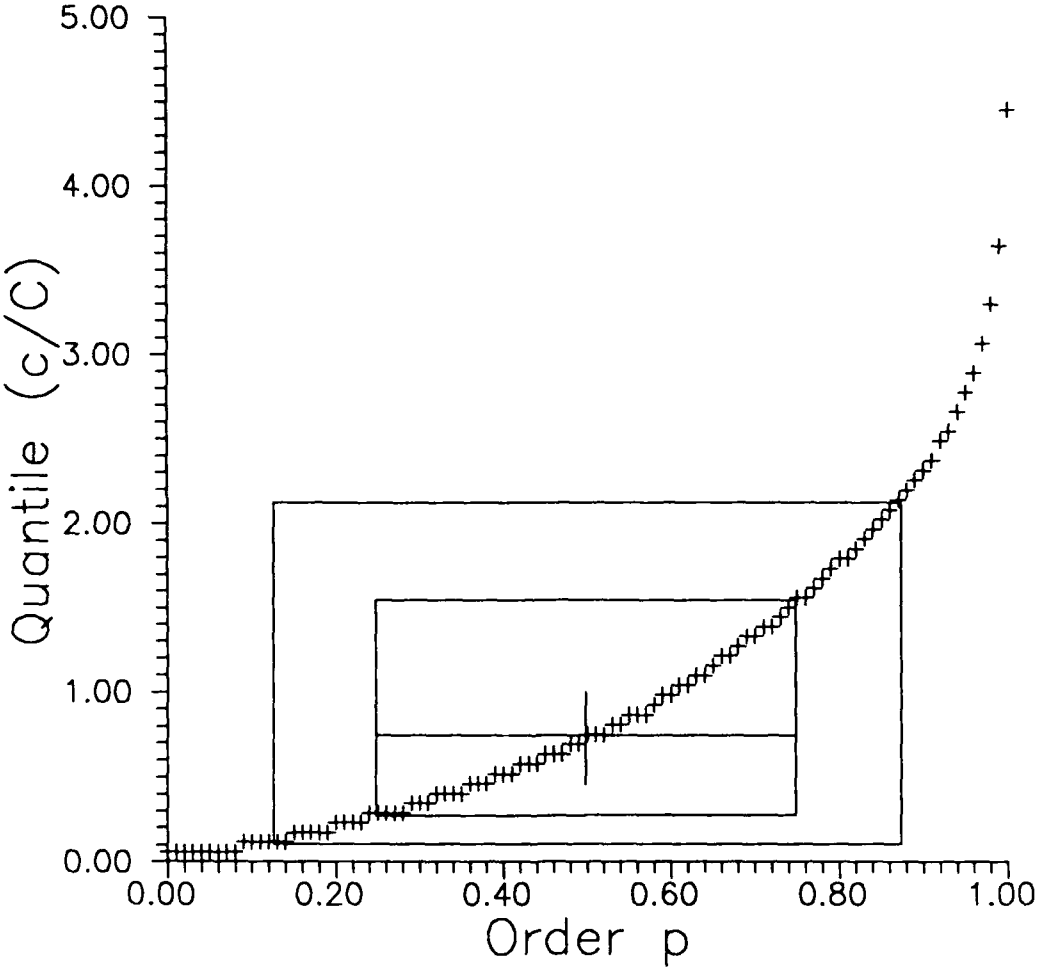

FIGURE 5

A quantile box plot for the conditional normalized time series from Sensor 4.

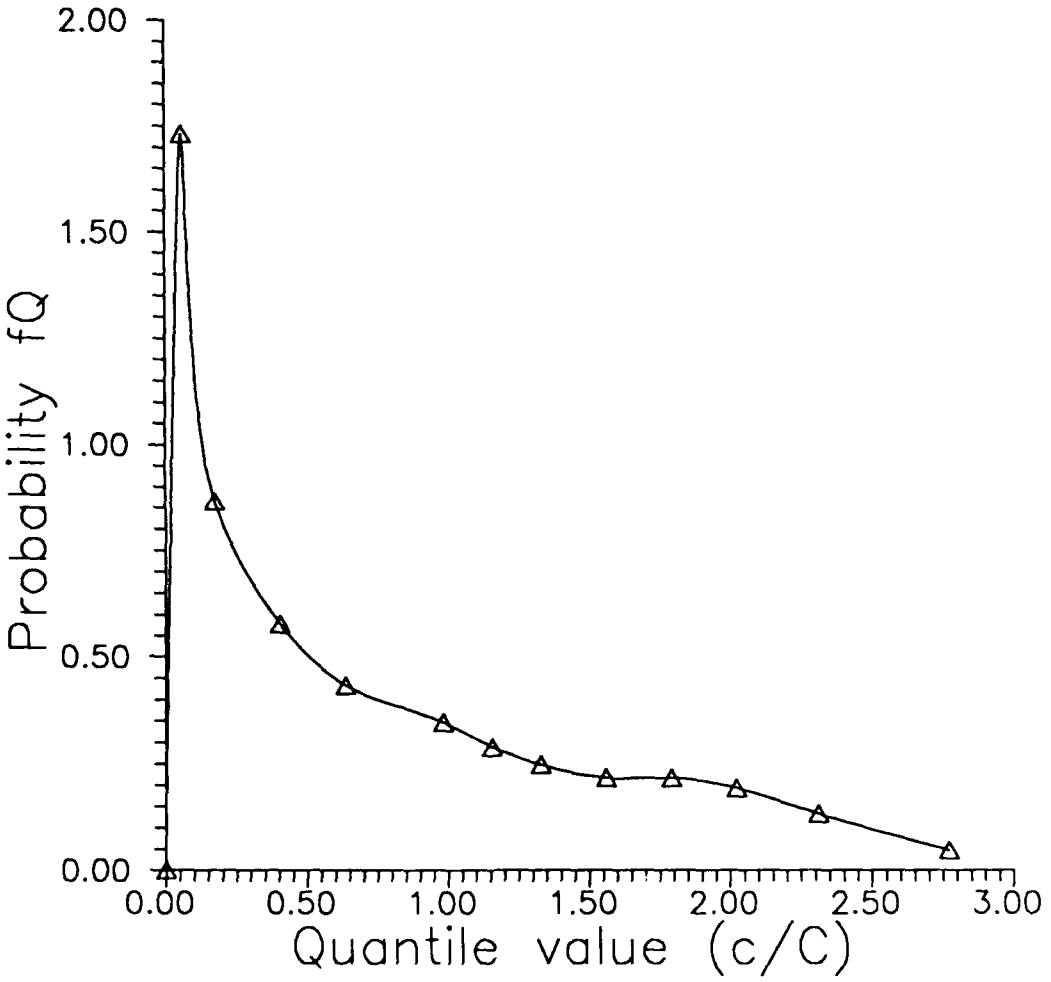

FIGURE 6

The sample density-quantile function for the conditional normalized time series from Sensor 4.

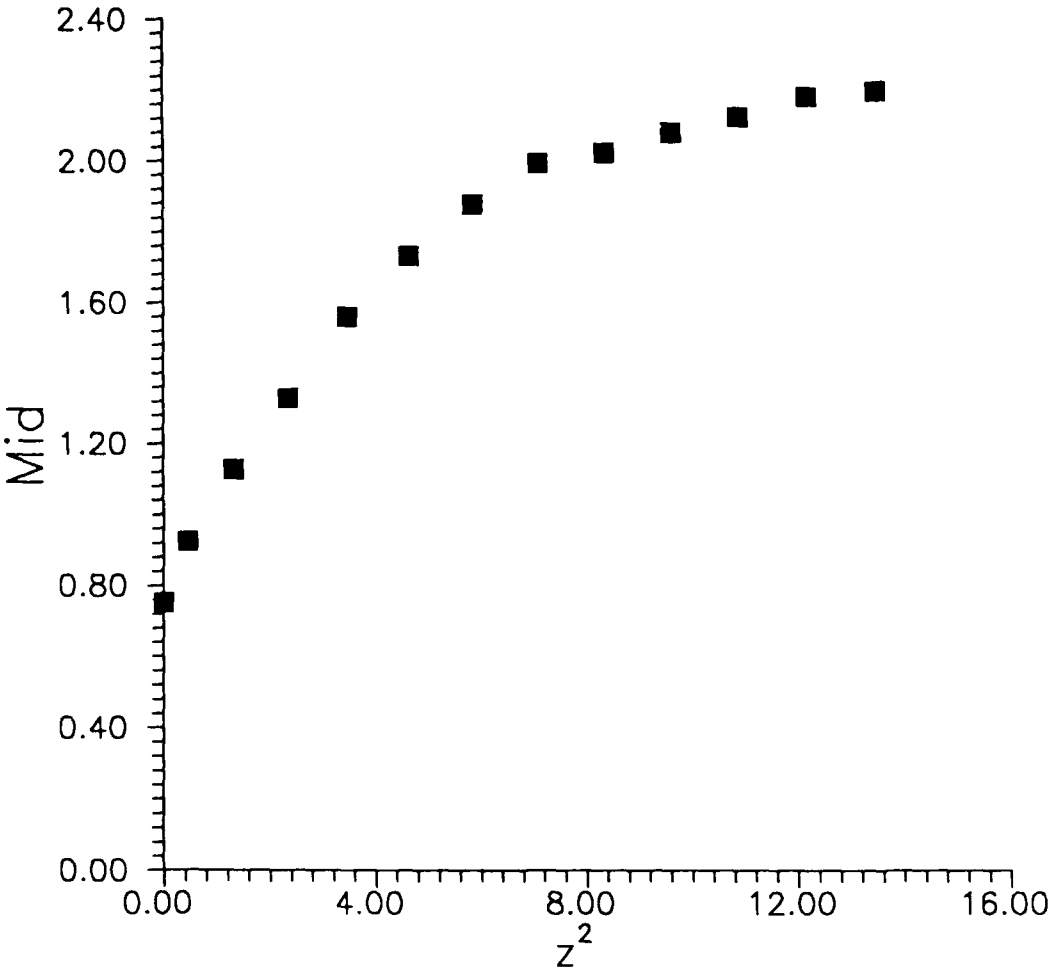

FIGURE 7

The mid-versus-z^2 plot for the conditional normalized time series from Sensor 4 illustrating the asymmetry in the data.

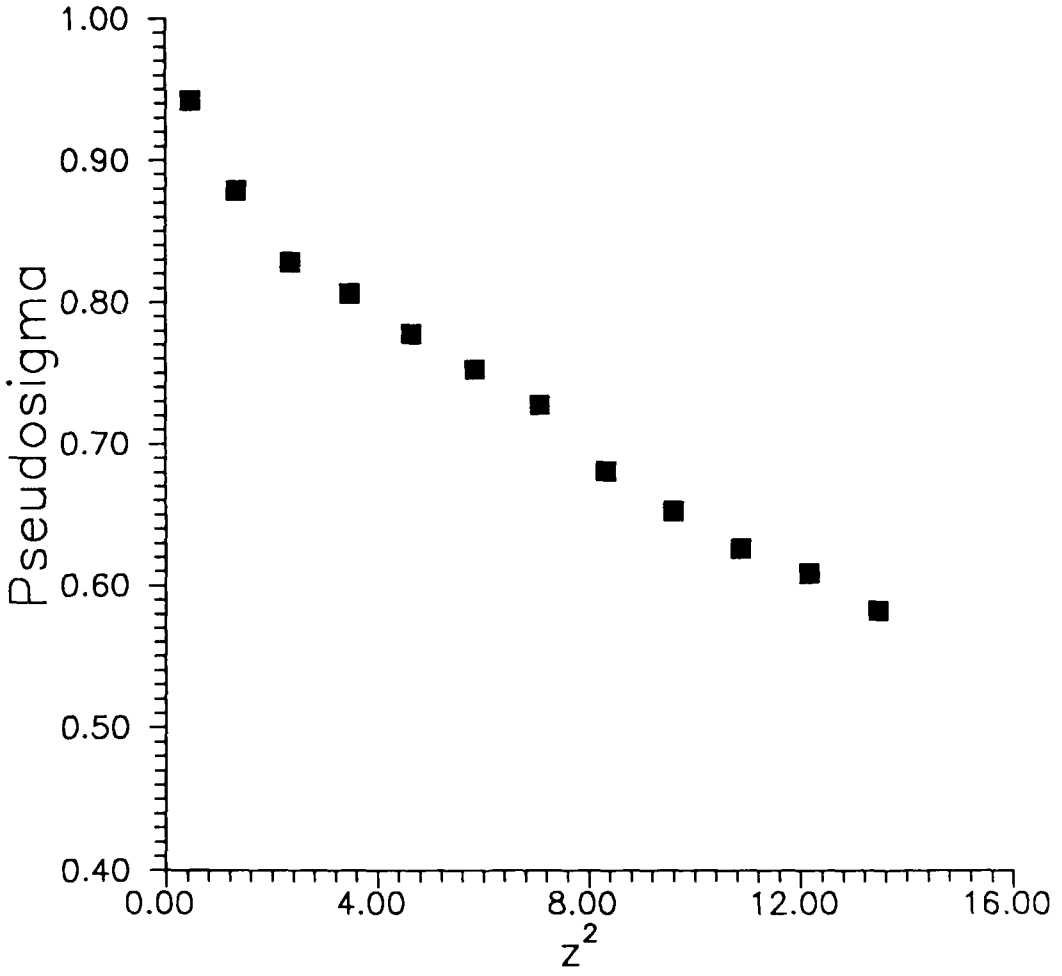

FIGURE 8

The pseudosigma-versus-z^2 plot for the conditional normalized time series from Sensor 4 illustrating the amount of elongation (i.e., the tailweight) present in the data.

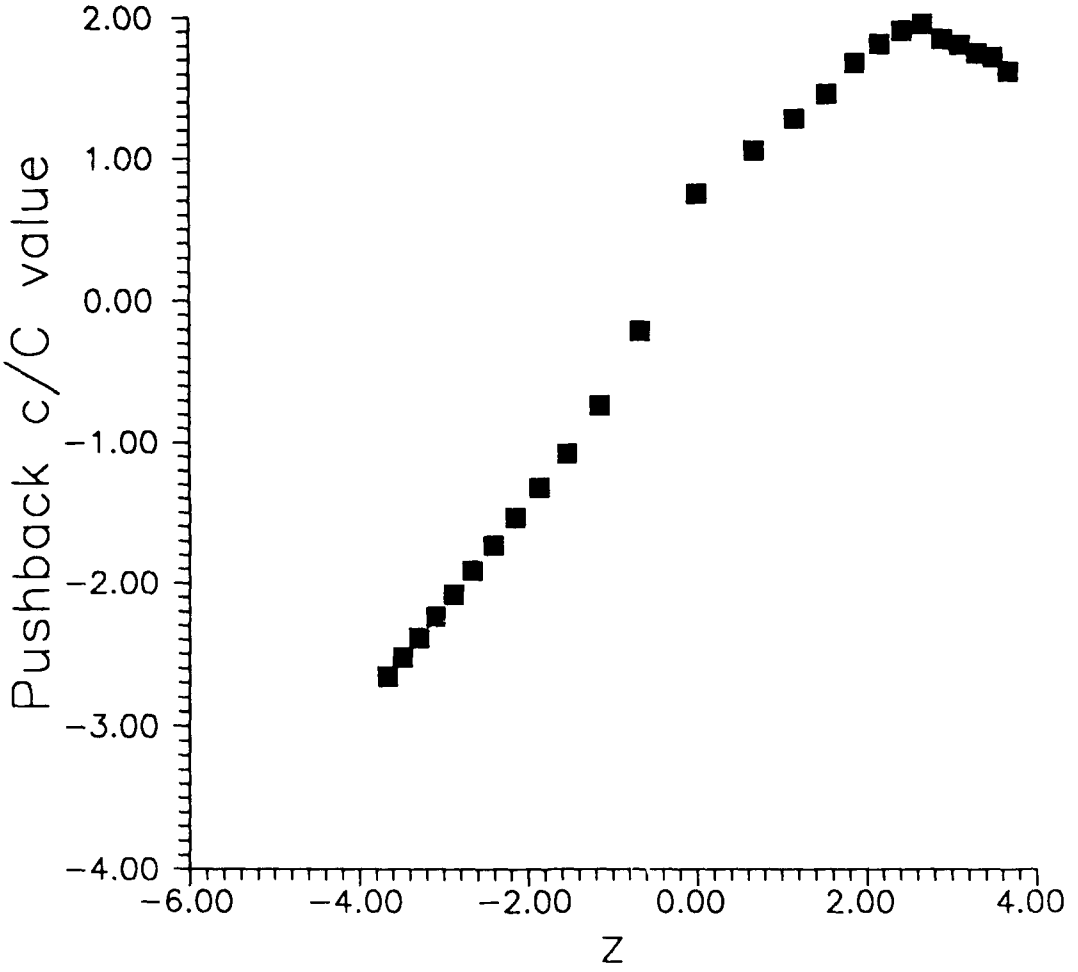

FIGURE 9

The pushback c/C values versus z for the conditional normalized time series from Sensor 4 revealing that the probability distribution for concentration is well approximated by three Gaussians with different scales.

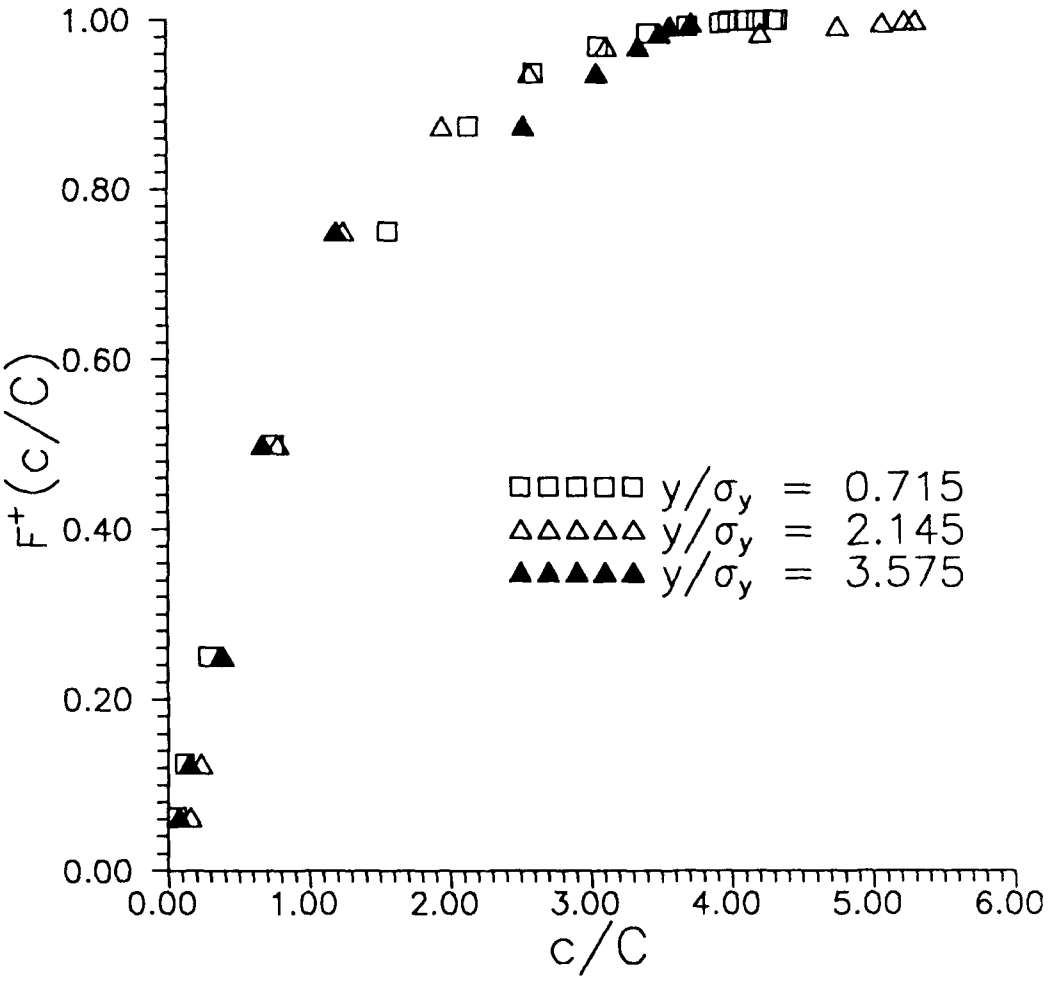

FIGURE 10

Conditional sample probability distribution functions obtained at three different lateral positions in the plume.

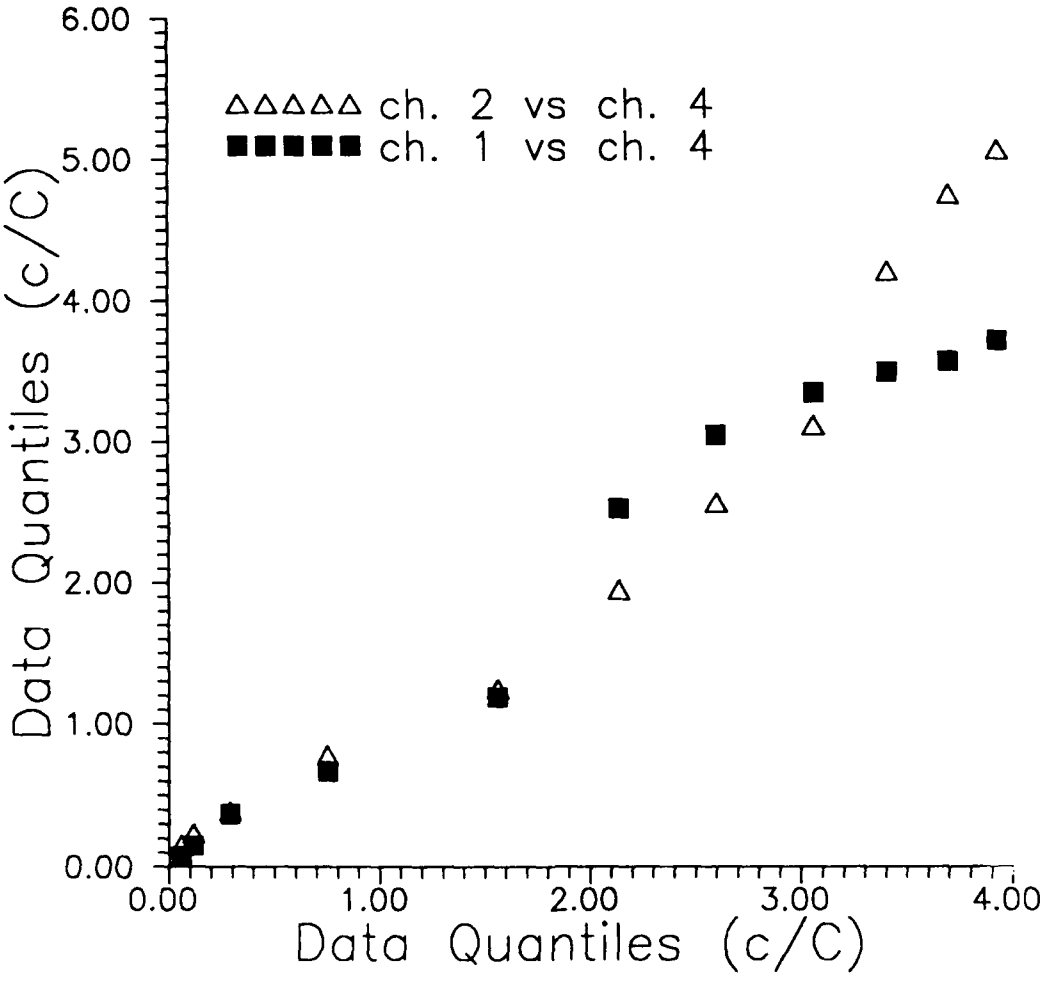

FIGURE 11

The Q-Q plot of the data quantiles of Sensors 1 and 2 versus the data quantiles of Sensor 4. The plot was constructed from the conditional normalized concentration data obtained from each of the sensors by censoring the periods of zero concentration and scaling the resulting data sequence by the conditional mean.

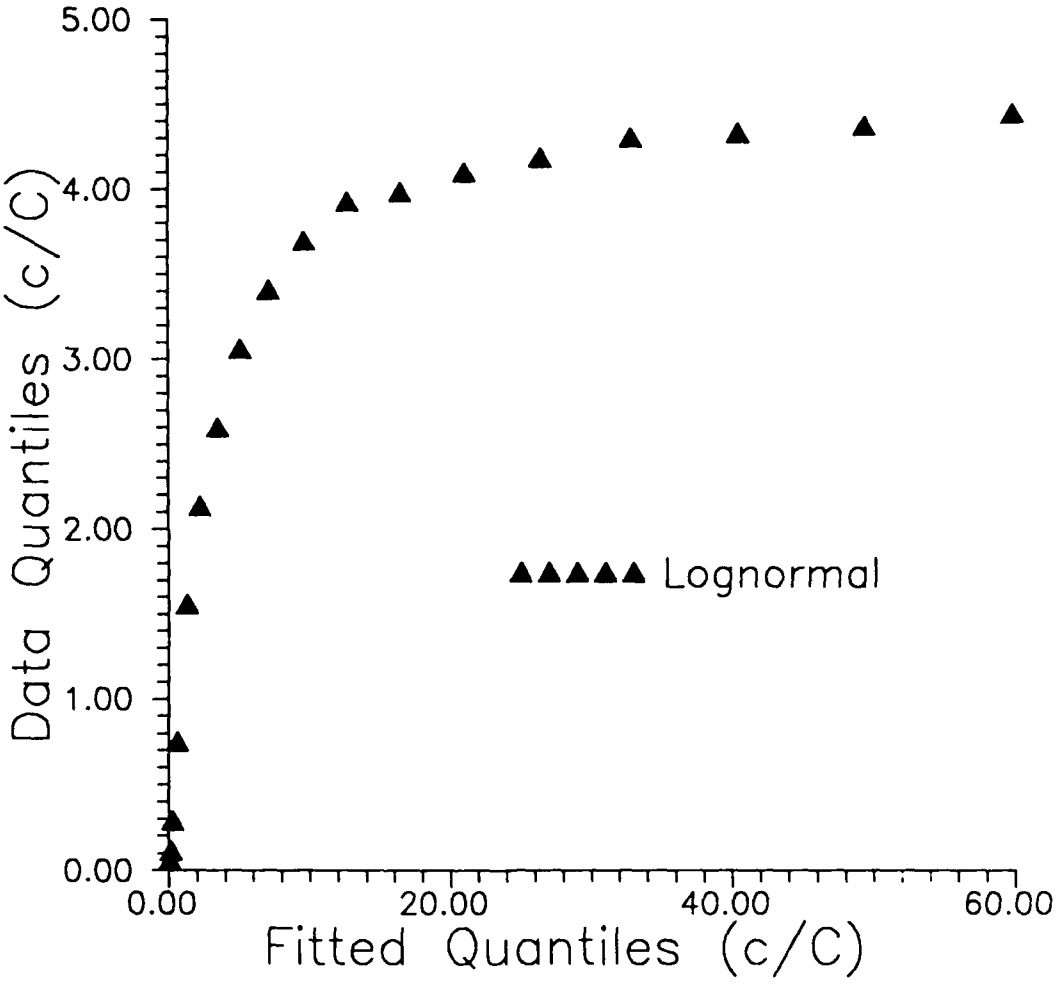

FIGURE 12

The Q-Q plot comparing the data quantiles of Sensor 4 with the associated quantiles of the fitted log-normal distribution.

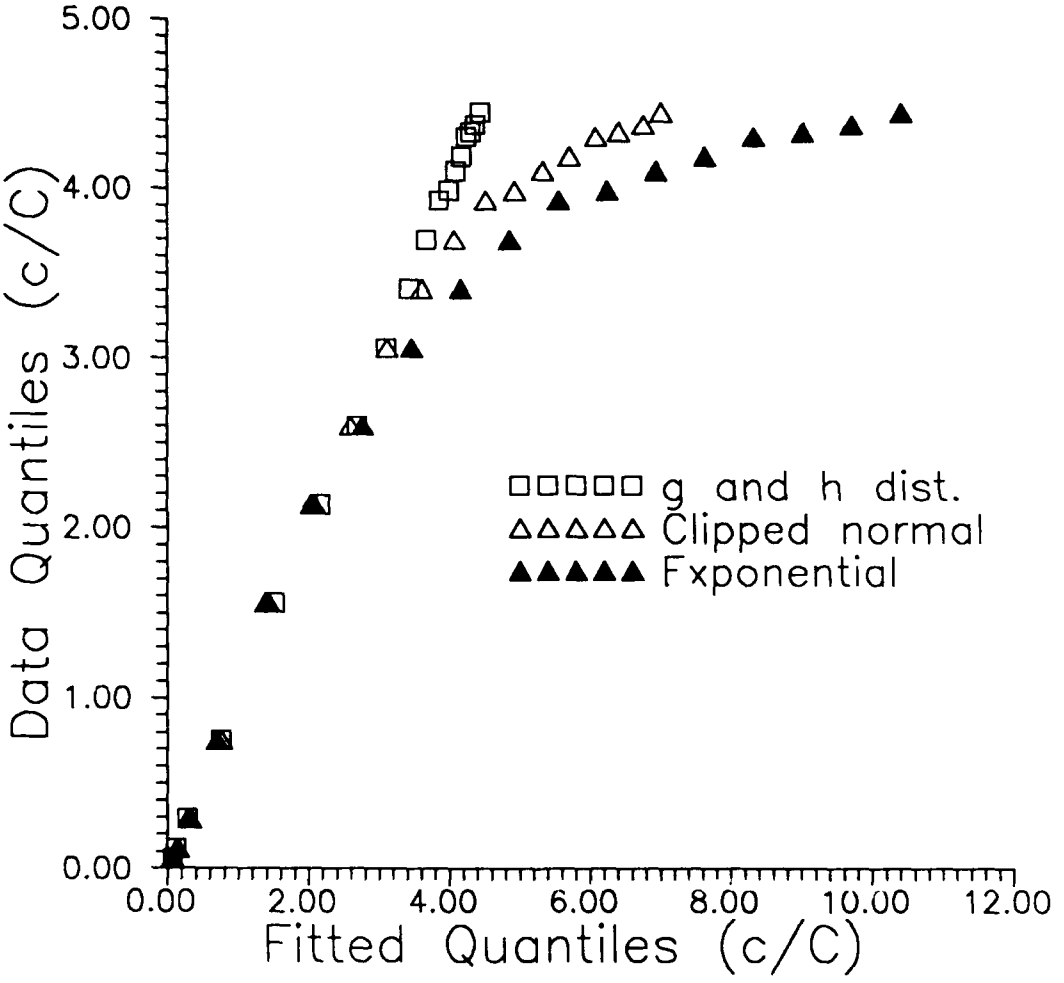

FIGURE 13

The Q-Q plot comparing the data quantiles of Sensor 4 with the associated quantiles of the fitted exponential, clipped-normal, and g and h distributions.

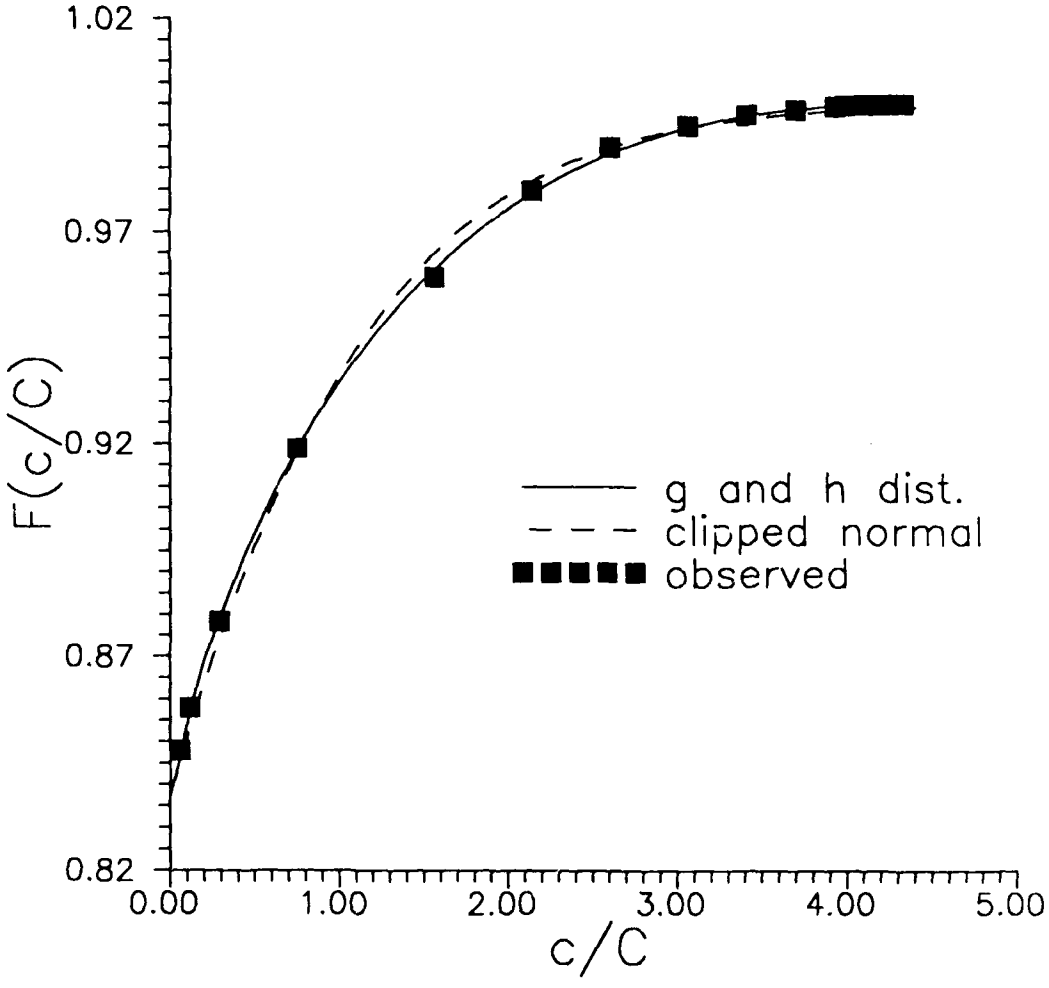

FIGURE 14

A comparison of the observed total probability distribution for Sensor 4 with the model total probability distributions obtained from the fitted clipped-normal and g and h distributions.

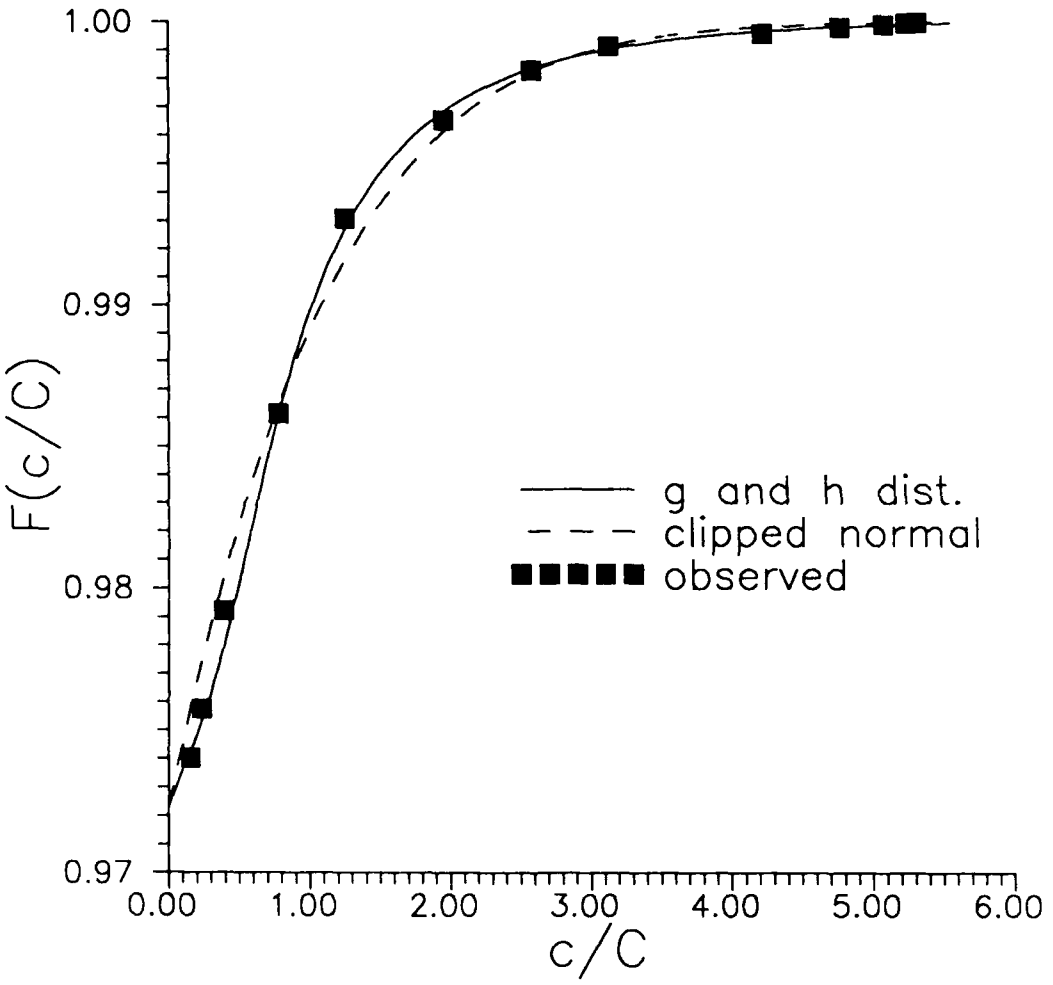

FIGURE 15

A comparison of the observed total probability distribution for Sensor 2 with the model total probability distributions obtained from the fitted clipped-normal and g and h distributions.

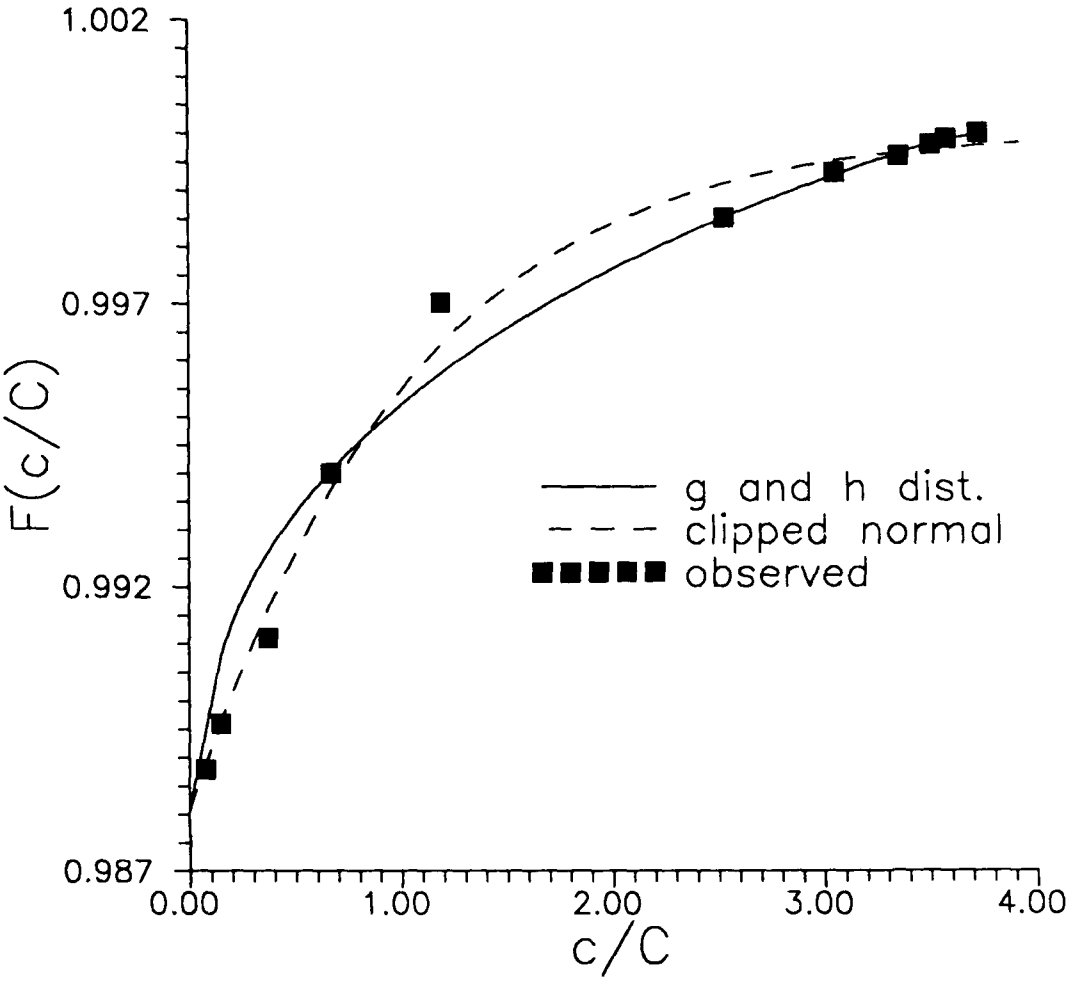

FIGURE 16

A comparison of the observed total probability distribution for Sensor 1 with the model total probability distributions obtained from the fitted clipped-normal and g and h distributions.

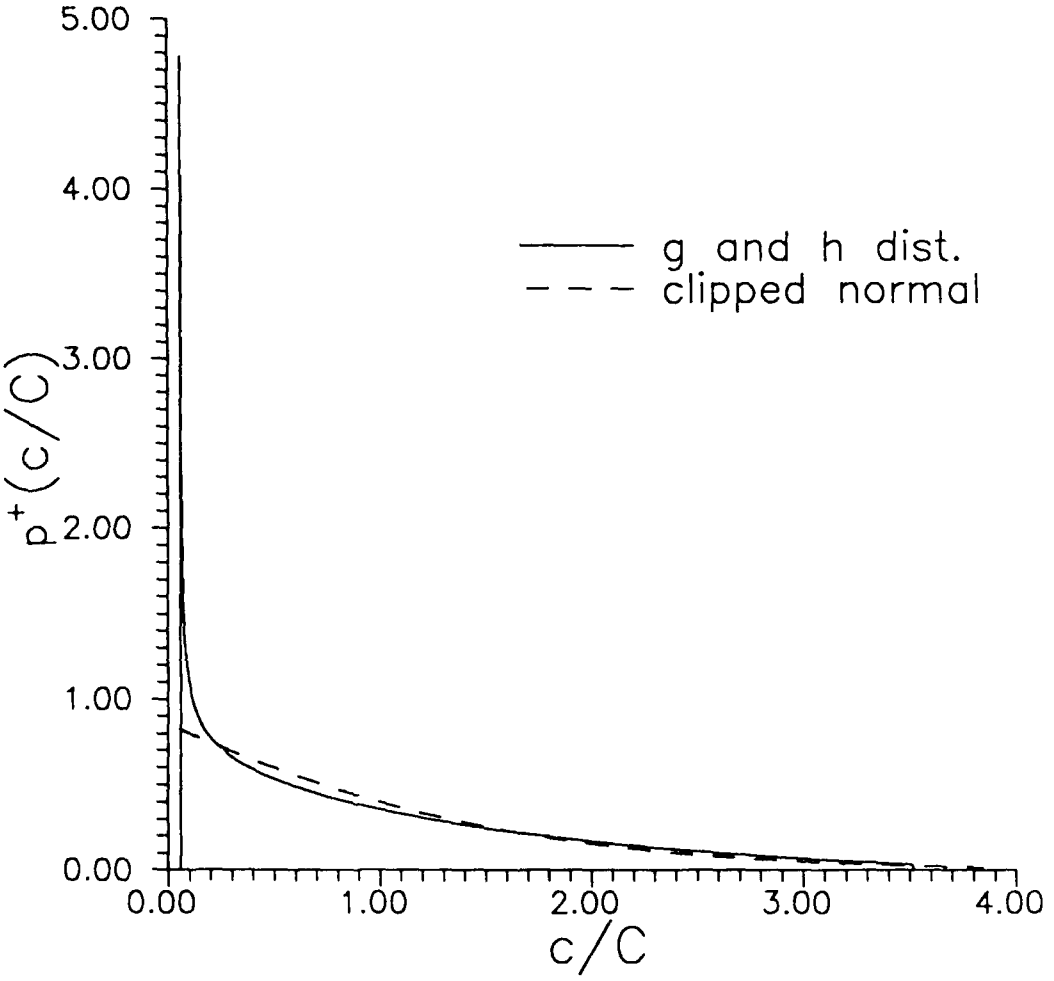

FIGURE 17

The fitted clipped-normal and g and h conditional probability density functions obtained for the conditional concentration data from Sensor 4 ($y/\sigma_y = 0.715$).

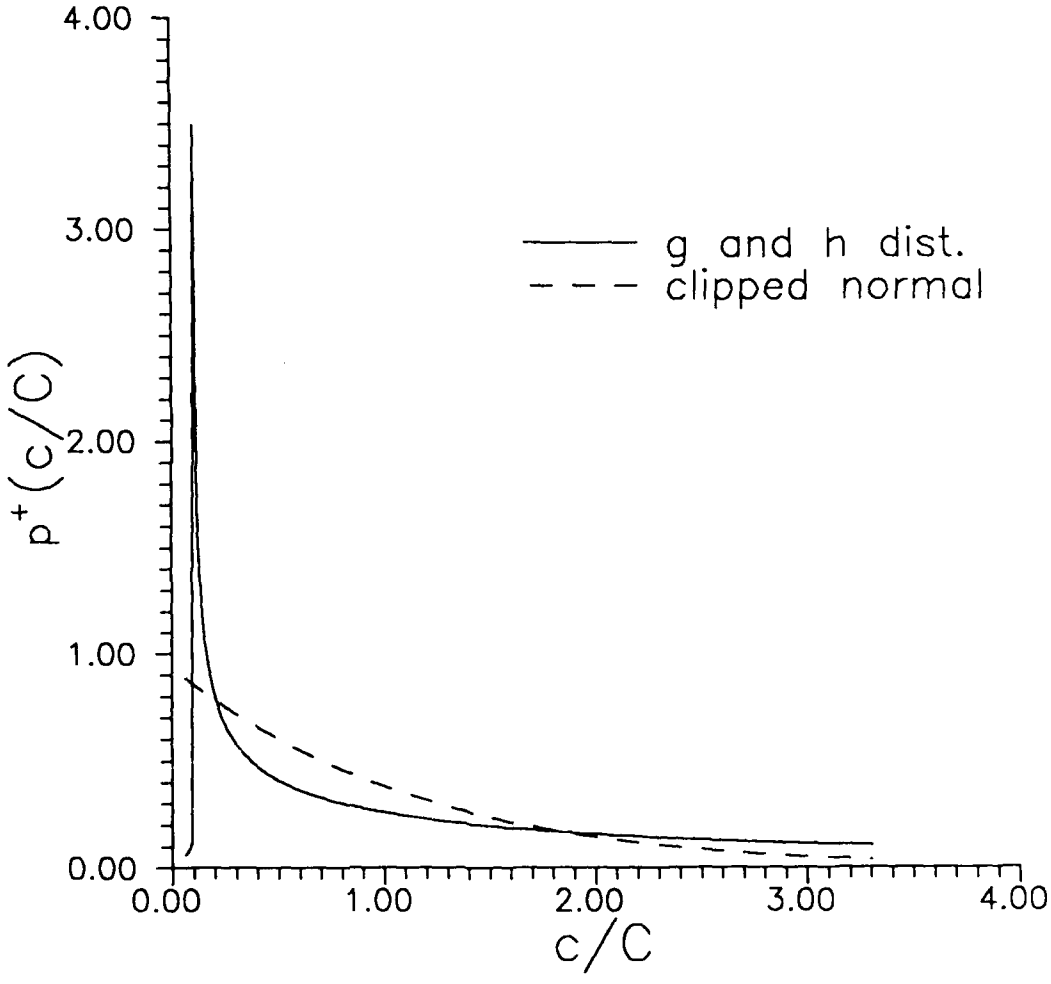

FIGURE 18

The fitted clipped-normal and g and h conditional probability density functions obtained for the conditional concentration data from Sensor 1 ($y/\sigma_y = 3.575$).

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence Research Establishment Suffield
Ralston, Alberta

2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

THE SHAPE OF THE PROBABILITY DENSITY FUNCTION OF SHORT-TERM CONCENTRATION FLUCTUATIONS OF PLUMES IN THE ATMOSPHERIC BOUNDARY LAYER (U)

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

YEE, Eugene C.

5. DATE OF PUBLICATION (month and year of publication of document)

July 1989

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

42

6b. NO. OF REFS (total cited in document)

18

6. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Suffield Report No. 525 (Final) (Jan, 1989 - June 1989)

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

PCN No. 051SP

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

SR 525

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(XX) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

UNLIMITED DISTRIBUTION

13 ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

The shape of the probability distribution of a set of high-resolution concentration fluctuation measurements from an ion plume is studied using order statistics and certain selected quantiles derived from them. A number of graphical techniques based on the order statistics are shown to be useful for the assessment of the symmetry and tailweight of the underlying distribution of concentration. These graphical techniques are applied, from both a descriptive and a computational point of view, to elucidate the underlying distributional shape of concentration and to assess the characterization efficacy of the probability distributions that have been proposed as models for concentration fluctuations. In this respect, a new probability distribution, namely, the g and h distribution, is introduced for describing concentration fluctuations and it is shown that this distribution is superior to the more commonly used models, namely, the log-normal, the exponential, and the clipped-normal distributions utilized by previous investigators. Except for the g and h distribution, it is found that none of the commonly used models for the concentration probability distribution is able to accurately characterize the extreme upper end of the concentration frequency distribution (i.e., the end of the distribution that is critical for the prediction of the probability of exposure to peak levels). However, the clipped-normal distribution is shown to provide a reasonably conservative model for the prediction of the exceedances of critical concentration levels. Finally, it is noted that the g and h distribution yields a bimodal form for the total probability density function for concentration whereas the clipped-normal distribution provides a unimodal form. It is shown that the bimodal form of the total concentration probability density function is consistent with both the data and certain theoretical results.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

CONCENTRATION FLUCTUATIONS
PROBABILITY DENSITY FUNCTION
ION TRACER
ORDER STATISTICS
EXPLORATORY DATA ANALYSIS
G- AND H- DISTRIBUTION
DISPERSION